

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 17, 2003, 18:19:29 ; Search time 54 Seconds
(without alignments)
824.721 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MEFSLLPRLPCNGAISAH.....FIRGVSPLYSGNSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	375	11 US-09-964-666-2	Sequence 2, Appli
2	2034	100.0	375	11 US-09-964-412-2	Sequence 2, Appli
3	2034	100.0	375	12 US-09-964-667-2	Sequence 2, Appli
4	2034	100.0	375	12 US-09-872-968-2	Sequence 2, Appli
5	2034	100.0	375	15 US-10-146-130-2	Sequence 2, Appli
6	2034	100.0	375	15 US-10-092-934-10	Sequence 10, Appli
7	2034	100.0	375	15 US-10-153-334-1	Sequence 1, Appli
8	2034	100.0	375	15 US-10-198-059-1	Sequence 1, Appli
9	2034	100.0	375	15 US-10-157-031-299	Sequence 299, App
10	2034	100.0	375	15 US-10-198-070-1	Sequence 1, Appli
11	734.5	36.1	361	11 US-09-995-494-107	Sequence 107, App
12	729	35.8	449	15 US-10-007-280A-140	Sequence 140, App
13	428.5	21.1	183	11 US-09-989-970-245	Sequence 245, App
14	382.5	18.8	673	15 US-10-157-031-291	Sequence 291, App
15	332	16.3	108	15 US-10-078-090-143	Sequence 143, App
16	316	15.5	152	15 US-10-198-070-112	Sequence 112, App

17	308	15.1	93	12 US-09-764-891-5337	Sequence 5337, Ap
18	308	15.1	93	15 US-10-205-428-486	Sequence 486, App
19	308	15.1	100	15 US-10-015-157A-187	Sequence 187, App
20	294.5	14.5	92	15 US-10-198-070-51	Sequence 51, Appl
21	281.5	13.8	115	11 US-09-995-494-79	Sequence 79, Appl
22	276	13.6	122	11 US-09-764-847-869	Sequence 869, Appl
23	276	13.6	122	15 US-10-092-154-869	Sequence 869, App
24	275	13.5	105	11 US-09-989-919-121	Sequence 121, App
25	273.5	13.4	118	15 US-10-198-070-30	Sequence 30, Appl
26	270	13.3	126	15 US-10-082-830-207	Sequence 207, App
27	264	13.0	91	15 US-10-001-857-198	Sequence 198, App
28	263	12.9	48	15 US-10-153-334-48	Sequence 48, Appl
29	262	12.9	118	15 US-10-198-070-116	Sequence 116, App
30	261	12.8	66	15 US-10-198-070-61	Sequence 61, Appl
31	260.5	12.8	109	11 US-09-989-919-105	Sequence 105, App
32	259	12.7	172	11 US-09-764-864-1058	Sequence 1058, Ap
33	259	12.7	232	15 US-10-198-070-34	Sequence 34, Appl
34	254.5	12.5	239	10 US-09-800-729-193	Sequence 193, App
35	254.5	12.5	239	15 US-10-198-070-113	Sequence 113, App
36	254.5	12.5	310	10 US-09-800-729-190	Sequence 190, App
37	254	12.5	147	12 US-09-764-891-3556	Sequence 3556, Ap
38	253	12.4	171	15 US-10-198-070-39	Sequence 39, Appl
39	251.5	12.4	91	15 US-10-011-585A-156	Sequence 156, App
40	251.5	12.4	144	15 US-10-011-585A-145	Sequence 145, App
41	251	12.3	776	15 US-10-020-079-24	Sequence 24, Appl
42	251	12.3	789	15 US-10-020-079-22	Sequence 22, Appl
43	251	12.3	863	15 US-10-020-079-32	Sequence 32, Appl
44	251	12.3	876	15 US-10-020-079-30	Sequence 30, Appl
45	251	12.3	889	15 US-10-020-079-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-964-666-2
; Sequence 2, Application US/09964666
; Patent No. US20020104108A1
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
Wands, Jack R.
TITLE OF INVENTION: Transgenic Animals and Cell Lines for
Screening Drugs Effective for the Treatment or Prevention
of Alzheimer's Disease
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSER: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
STREET: 1100 New York Ave., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/964,666
FILING DATE: 28-Sep-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0609.4370000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-666-2

Query Match 100.0%; Score 2034; DB 11; Length 375;
Best Local Similarity 100.0%; Pred. No. 3,9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFLLPRLCNGAISHNRRLRPGSSDSPASAPVAGITGMCTHARLLIYFLVEMEF 60
DB 1 MEFLLPRLCNGAISHNRRLRPGSSDSPASAPVAGITGMCTHARLLIYFLVEMEF 60
QY 61 LHVQAGLELPTSDPSVASQSARVYRTGHHARCLANFCGRNRVSLMCPSPSELKOST 120
DB 61 LHVQAGLELPTSDPSVASQSARVYRTGHHARCLANFCGRNRVSLMCPSPSELKOST 120
QY 121 CLSLPKCWDYRAAVPGLEFLFLRHRCPTLTODEVQCDHSSLQSPTEIKHPASASQ 180
DB 121 CLSLPKCWDYRAAVPGLEFLFLRHRCPTLTODEVQCDHSSLQSPTEIKHPASASQ 180
QY 181 VAGTRDMHHTWLIFFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPFGFKLFSCPSLSS 240
DB 181 VAGTRDMHHTWLIFFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPFGFKLFSCPSLSS 240
QY 241 WDYRRPRLANFVFLVEMGFTMFARLLISGCDLPASASQAGITGVSHHARLIENFC 300
DB 241 WDYRRPRLANFVFLVEMGFTMFARLLISGCDLPASASQAGITGVSHHARLIENFC 300
QY 301 LFEHESHVTOAGVQWPNLGSLOPLPGLKRFSCSLPSSWDYGHLPHPANFCIFIRG 360
DB 301 LFEHESHVTOAGVQWPNLGSLOPLPGLKRFSCSLPSSWDYGHLPHPANFCIFIRG 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375

RESULT 2

US-09-964-412-2
Sequence 2, Application US/09964412
Patent No. US20020129391A1

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
Wands, Jack R.

TITLE OF INVENTION: Transgenic Animals and Cell Lines for
Screening Drugs Effective for the Treatment or Prevention
of Alzheimer's Disease

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS: Kessler, Goldstein & Fox, P.L.L.C.

ADDRESS: Sterne, Kessler, Goldstein & Fox, P.L.L.C.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/964,412

FILING DATE: 28-Sep-2001

CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0609,4370000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-412-2

Query Match 100.0%; Score 2034; DB 11; Length 375;
Best Local Similarity 100.0%; Pred. No. 3,9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFLLPRLCNGAISHNRRLRPGSSDSPASAPVAGITGMCTHARLLIYFLVEMEF 60
DB 1 MEFLLPRLCNGAISHNRRLRPGSSDSPASAPVAGITGMCTHARLLIYFLVEMEF 60
QY 61 LHVQAGLELPTSDPSVASQSARVYRTGHHARCLANFCGRNRVSLMCPSPSELKOST 120
DB 61 LHVQAGLELPTSDPSVASQSARVYRTGHHARCLANFCGRNRVSLMCPSPSELKOST 120
QY 121 CLSLPKCWDYRAAVPGLEFLFLRHRCPTLTODEVQCDHSSLQSPTEIKHPASASQ 180
DB 121 CLSLPKCWDYRAAVPGLEFLFLRHRCPTLTODEVQCDHSSLQSPTEIKHPASASQ 180
QY 181 VAGTRDMHHTWLIFFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPFGFKLFSCPSLSS 240
DB 181 VAGTRDMHHTWLIFFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPFGFKLFSCPSLSS 240
QY 241 WDYRRPRLANFVFLVEMGFTMFARLLISGCDLPASASQAGITGVSHHARLIENFC 300
DB 241 WDYRRPRLANFVFLVEMGFTMFARLLISGCDLPASASQAGITGVSHHARLIENFC 300
QY 301 LFEHESHVTOAGVQWPNLGSLOPLPGLKRFSCSLPSSWDYGHLPHPANFCIFIRG 360
DB 301 LFEHESHVTOAGVQWPNLGSLOPLPGLKRFSCSLPSSWDYGHLPHPANFCIFIRG 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375

RESULT 3

US-09-964-667-2
Sequence 2, Application US/09964667
Publication No. US2003003621A1

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
Wands, Jack R.

TITLE OF INVENTION: Transgenic Animals and Cell Lines for
Screening Drugs Effective for the Treatment or Preventio
of Alzheimer's Disease

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS: Kessler, Goldstein & Fox, P.L.L.C.

ADDRESS: Sterne, Kessler, Goldstein & Fox, P.L.L.C.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/964,667

FILING DATE: 28-Sep-2001

CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0609,4370000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-667-2

Query Match 100.0%; Score 2034; DB 12; Length 375;
Best Local Similarity 100.0%; Pred. No. 3.9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLCEGNGAISAHRNRLPGSSDPSASASPVAGITGMCTHARLILYFFLEVEMF 60
DB 1 MEFSLLPRLCEGNGAISAHRNRLPGSSDPSASASPVAGITGMCTHARLILYFFLEVEMF 60
QY 61 LHVGAAGLELPTSDPSVASASOSARYRTGHHARLCLANFCGRNRVSLMCPSPSPELKOST 120
DB 61 LHVGAAGLELPTSDPSVASASOSARYRTGHHARLCLANFCGRNRVSLMCPSPSPELKOST 120
QY 121 CUSLPKCMYRRAAAYPGLEFLFELRHRCPTLTODEVQWCHSSLOPSTPEIKHPPASASQ 180
DB 121 CUSLPKCMYRRAAAYPGLEFLFELRHRCPTLTODEVQWCHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTDMHHYTWLIFIFINFLROSLSNVTQAGVQWRNLSLOPLPPGFKLSPSCSLSS 240
DB 181 VAGTDMHHYTWLIFIFINFLROSLSNVTQAGVQWRNLSLOPLPPGFKLSPSCSLSS 240
QY 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
DB 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
QY 301 LFMESHSTYQAGVQWPNLGSLOPLPPGKRRSCSLSPSSMDYGHLPHPANFCIFIRGG 360
DB 301 LFMESHSTYQAGVQWPNLGSLOPLPPGKRRSCSLSPSSMDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGMSQTPDLR 375
DB 361 VSPYLSGMSQTPDLR 375

RESULT 4
US-09-872-968-2

Sequence 2, Application US/09872968
Publication No. US20030050262A1
GENERAL INFORMATION:
APPLICANT: Wands, Jack R
APPLICANT: de la Monte, Suzanne M
TITLE OF INVENTION: Inhibition of Neurodegeneration
FILE REFERENCE: 21486-047
CURRENT APPLICATION NUMBER: US/09/872,968
CURRENT FILING DATE: 2001-06-01
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2

LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
US-09-968-2

Query Match 100.0%; Score 2034; DB 12; Length 375;
Best Local Similarity 100.0%; Pred. No. 3.9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLCEGNGAISAHRNRLPGSSDPSASASPVAGITGMCTHARLILYFFLEVEMF 60
DB 1 MEFSLLPRLCEGNGAISAHRNRLPGSSDPSASASPVAGITGMCTHARLILYFFLEVEMF 60
QY 61 LHVGAAGLELPTSDPSVASASOSARYRTGHHARLCLANFCGRNRVSLMCPSPSPELKOST 120
DB 61 LHVGAAGLELPTSDPSVASASOSARYRTGHHARLCLANFCGRNRVSLMCPSPSPELKOST 120

QY 121 CUSLPKCMYRRAAAYPGLEFLFELRHRCPTLTODEVQWCHSSLOPSTPEIKHPPASASQ 180
DB 121 CUSLPKCMYRRAAAYPGLEFLFELRHRCPTLTODEVQWCHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTDMHHYTWLIFIFINFLROSLSNVTQAGVQWRNLSLOPLPPGFKLSPSCSLSS 240
DB 181 VAGTDMHHYTWLIFIFINFLROSLSNVTQAGVQWRNLSLOPLPPGFKLSPSCSLSS 240
QY 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
DB 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
QY 301 LFMESHSTYQAGVQWPNLGSLOPLPPGKRRSCSLSPSSMDYGHLPHPANFCIFIRGG 360
DB 301 LFMESHSTYQAGVQWPNLGSLOPLPPGKRRSCSLSPSSMDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGMSQTPDLR 375
DB 361 VSPYLSGMSQTPDLR 375

RESULT 5

US-10-146-130-2
Sequence 2, Application US/10146130
Publication No. US20030004107A1
GENERAL INFORMATION:

APPLICANT: AVERBACK, PAUL
TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF
FILE REFERENCE: 59003.000007
CURRENT APPLICATION NUMBER: US/10/146,130
CURRENT FILING DATE: 2002-08-06
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
US-10-146-130-2

Query Match 100.0%; Score 2034; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 3.9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLCEGNGAISAHRNRLPGSSDPSASASPVAGITGMCTHARLILYFFLEVEMF 60
DB 1 MEFSLLPRLCEGNGAISAHRNRLPGSSDPSASASPVAGITGMCTHARLILYFFLEVEMF 60
QY 61 LHVGAAGLELPTSDPSVASASOSARYRTGHHARLCLANFCGRNRVSLMCPSPSPELKOST 120
DB 61 LHVGAAGLELPTSDPSVASASOSARYRTGHHARLCLANFCGRNRVSLMCPSPSPELKOST 120
QY 121 CUSLPKCMYRRAAAYPGLEFLFELRHRCPTLTODEVQWCHSSLOPSTPEIKHPPASASQ 180
DB 121 CUSLPKCMYRRAAAYPGLEFLFELRHRCPTLTODEVQWCHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTDMHHYTWLIFIFINFLROSLSNVTQAGVQWRNLSLOPLPPGFKLSPSCSLSS 240
DB 181 VAGTDMHHYTWLIFIFINFLROSLSNVTQAGVQWRNLSLOPLPPGFKLSPSCSLSS 240
QY 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
DB 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
QY 301 LFMESHSTYQAGVQWPNLGSLOPLPPGKRRSCSLSPSSMDYGHLPHPANFCIFIRGG 360
DB 301 LFMESHSTYQAGVQWPNLGSLOPLPPGKRRSCSLSPSSMDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGMSQTPDLR 375
DB 361 VSPYLSGMSQTPDLR 375

```
RESULT 6
US-10-092-934-10
; Sequence 10, Application US/10092934
; Publication No. US20030054990A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHODS OF USING NEURAL THREAD PROTEINS TO TREAT TUMORS
; TITLE OF INVENTION: AND CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 018792-0199
; CURRENT APPLICATION NUMBER: US/10/092,934
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/273,957
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-934-10
```

```
Query Match      100.0%; Score 2034; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 3.9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MESSLLPRLCNGAISAHNRNLRPGSSDPSPASAPVAGITGMCTHARLLIYFELVMEF 60
DB 1 MESSLLPRLCNGAISAHNRNLRPGSSDPSPASAPVAGITGMCTHARLLIYFELVMEF 60
QY 61 LHVGAGLELPTSDPPSVASASQARYRTGHHARCLANFCGRNRYSLMCPMSPELKOST 120
DB 61 LHVGAGLELPTSDPPSVASASQARYRTGHHARCLANFCGRNRYSLMCPMSPELKOST 120
QY 121 CLSLPKCWDYRAAVPGFLIFELFRLHRCPTLTQDEVOMCHSSLOPSTPEIKHPASASQ 180
DB 121 CLSLPKCWDYRAAVPGFLIFELFRLHRCPTLTQDEVOMCHSSLOPSTPEIKHPASASQ 180
QY 121 VAGTRDMHHTWLIFFIFENFLRQSLNSVTQAGVQWNLGSLQPLPGFKLFSCPSLSS 240
DB 121 VAGTRDMHHTWLIFFIFENFLRQSLNSVTQAGVQWNLGSLQPLPGFKLFSCPSLSS 240
QY 241 WDYRRPRLANFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIENFC 300
DB 241 WDYRRPRLANFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIENFC 300
QY 301 LFEMESHVSVOAGVQWPNLGLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVSVOAGVQWPNLGLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGMSQTPDLR 375
DB 361 VSPYLSGMSQTPDLR 375
```

```
RESULT 7
US-10-153-334-1
; Sequence 1, Application US/10153334
; Publication No. US20030096350A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003-000006
; CURRENT APPLICATION NUMBER: US/10/153,334
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,156
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
```

```
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-334-1
```

```
Query Match      100.0%; Score 2034; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 3.9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MESSLLPRLCNGAISAHNRNLRPGSSDPSPASAPVAGITGMCTHARLLIYFELVMEF 60
DB 1 MESSLLPRLCNGAISAHNRNLRPGSSDPSPASAPVAGITGMCTHARLLIYFELVMEF 60
QY 61 LHVGAGLELPTSDPPSVASASQARYRTGHHARCLANFCGRNRYSLMCPMSPELKOST 120
DB 61 LHVGAGLELPTSDPPSVASASQARYRTGHHARCLANFCGRNRYSLMCPMSPELKOST 120
QY 121 CLSLPKCWDYRAAVPGFLIFELFRLHRCPTLTQDEVOMCHSSLOPSTPEIKHPASASQ 180
DB 121 CLSLPKCWDYRAAVPGFLIFELFRLHRCPTLTQDEVOMCHSSLOPSTPEIKHPASASQ 180
QY 121 VAGTRDMHHTWLIFFIFENFLRQSLNSVTQAGVQWNLGSLQPLPGFKLFSCPSLSS 240
DB 121 VAGTRDMHHTWLIFFIFENFLRQSLNSVTQAGVQWNLGSLQPLPGFKLFSCPSLSS 240
QY 241 WDYRRPRLANFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIENFC 300
DB 241 WDYRRPRLANFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIENFC 300
QY 301 LFEMESHVSVOAGVQWPNLGLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVSVOAGVQWPNLGLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGMSQTPDLR 375
DB 361 VSPYLSGMSQTPDLR 375
```

```
RESULT 8
US-10-198-069-1
; Sequence 1, Application US/10198069
; Publication No. US20030096756A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003-000009
; CURRENT APPLICATION NUMBER: US/10/198,069
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-069-1
```

```
Query Match      100.0%; Score 2034; DB 15; Length 375;
Best Local Similarity 100.0%; Pred. No. 3.9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MESSLLPRLCNGAISAHNRNLRPGSSDPSPASAPVAGITGMCTHARLLIYFELVMEF 60
DB 1 MESSLLPRLCNGAISAHNRNLRPGSSDPSPASAPVAGITGMCTHARLLIYFELVMEF 60
QY 61 LHVGAGLELPTSDPPSVASASQARYRTGHHARCLANFCGRNRYSLMCPMSPELKOST 120
```

Db 61 LHVGOAGLEPTSDPSVASASARYRTGHNARLCLANCGNRNRYSLMCPSPPELKOST 120
121 CUSLPRKCDYRRAAVPGLEFLFLLHRCPTLTODEVOWCDHSSLOPSTPEIKHPPASASQ 180
121 CUSLPRKCDYRRAAVPGLEFLFLLHRCPTLTODEVOWCDHSSLOPSTPEIKHPPASASQ 180
181 VAGTDMHHYTWLIFIFNFRSLNSTVTOAGVOMRNIGSLQPLPFGKFLSCPSLSS 240
181 VAGTDMHHYTWLIFIFNFRSLNSTVTOAGVOMRNIGSLQPLPFGKFLSCPSLSS 240
241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
301 LFEEMSHSVTOAGVOMPNGSLQPLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGC 360
301 LFEEMSHSVTOAGVOMPNGSLQPLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGC 360
361 VSPYLSGWSQTPDLR 375
361 VSPYLSGWSQTPDLR 375

RESULT 9

US-10-157-031-299
Sequence 299, Application US/10157031
Publication No. US20030108890A1

GENERAL INFORMATION:

APPLICANT: Baranova, A. V.
APPLICANT: Yankovsky, N. K.
APPLICANT: Kozlov, A. P.
APPLICANT: Lobashev, A. V.
APPLICANT: Krakovskaya, L. L.
TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
FILE REFERENCE: 2760-103
CURRENT APPLICATION NUMBER: US/10/157,031
-CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 415
SOFTWARE: PatentIn version 3.1
SEQ ID NO 299
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
US-10-157-031-299

Query Match 100.0%, Score 2034, DB 15, Length 375;
Best Local Similarity 100.0%, Pred. No. 3.9e-184;

Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLCEGCAISAHNRRLRPGSSDPSASASPVAGITGMCTHARLLIFLVEMEF 60
Db 1 MEFSLLPRLCEGCAISAHNRRLRPGSSDPSASASPVAGITGMCTHARLLIFLVEMEF 60
QY 61 LHVGOAGLEPTSDPSVASASARYRTGHNARLCLANCGNRNRYSLMCPSPPELKOST 120
Db 61 LHVGOAGLEPTSDPSVASASARYRTGHNARLCLANCGNRNRYSLMCPSPPELKOST 120
QY 121 CUSLPRKCDYRRAAVPGLEFLFLLHRCPTLTODEVOWCDHSSLOPSTPEIKHPPASASQ 180
Db 121 CUSLPRKCDYRRAAVPGLEFLFLLHRCPTLTODEVOWCDHSSLOPSTPEIKHPPASASQ 180
181 VAGTDMHHYTWLIFIFNFRSLNSTVTOAGVOMRNIGSLQPLPFGKFLSCPSLSS 240
181 VAGTDMHHYTWLIFIFNFRSLNSTVTOAGVOMRNIGSLQPLPFGKFLSCPSLSS 240
241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
QY 301 LFEEMSHSVTOAGVOMPNGSLQPLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGC 360
Db 301 LFEEMSHSVTOAGVOMPNGSLQPLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGC 360

QY 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 10

US-10-198-070-1
Sequence 1, Application US/10198070
Publication No. US20030109437A1

GENERAL INFORMATION:

APPLICANT: AVERBACK, PAUL
APPLICANT: GEMMELL, JACK
TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF CELLS
FILE REFERENCE: 59003.000008
CURRENT APPLICATION NUMBER: US/10/198,070
-CURRENT FILING DATE: 2002-07-19
PRIOR APPLICATION NUMBER: 60/306,161
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/306,150
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/331,477
PRIOR FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 125
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
US-10-198-070-1

Query Match 100.0%, Score 2034, DB 15, Length 375;
Best Local Similarity 100.0%, Pred. No. 3.9e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLCEGCAISAHNRRLRPGSSDPSASASPVAGITGMCTHARLLIFLVEMEF 60
Db 1 MEFSLLPRLCEGCAISAHNRRLRPGSSDPSASASPVAGITGMCTHARLLIFLVEMEF 60
QY 61 LHVGOAGLEPTSDPSVASASARYRTGHNARLCLANCGNRNRYSLMCPSPPELKOST 120
Db 61 LHVGOAGLEPTSDPSVASASARYRTGHNARLCLANCGNRNRYSLMCPSPPELKOST 120
QY 121 CUSLPRKCDYRRAAVPGLEFLFLLHRCPTLTODEVOWCDHSSLOPSTPEIKHPPASASQ 180
Db 121 CUSLPRKCDYRRAAVPGLEFLFLLHRCPTLTODEVOWCDHSSLOPSTPEIKHPPASASQ 180
181 VAGTDMHHYTWLIFIFNFRSLNSTVTOAGVOMRNIGSLQPLPFGKFLSCPSLSS 240
181 VAGTDMHHYTWLIFIFNFRSLNSTVTOAGVOMRNIGSLQPLPFGKFLSCPSLSS 240
241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
QY 301 LFEEMSHSVTOAGVOMPNGSLQPLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGC 360
Db 301 LFEEMSHSVTOAGVOMPNGSLQPLPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGC 360
QY 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 11

US-09-995-494-107
Sequence 107, Application US/09995494
Patent No. US20020127578A1

GENERAL INFORMATION:

APPLICANT: Salceda, Susana
APPLICANT: Macina, Roberto

```

: APPLICANT: Recipon, Herve
: APPLICANT: Cafferey, Robert
: APPLICANT: Ali, Shujath
: APPLICANT: Sun, Yongming
: APPLICANT: Liu, Chenghua
: APPLICANT: Chen, Sei-Yu
: TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
: FILE REFERENCE: DEX-0293
: CURRENT APPLICATION NUMBER: US/09/995,494
: CURRENT FILING DATE: 2001-11-27
: PRIOR APPLICATION NUMBER: 60/253,176
: NUMBER OF SEQ ID NOS: 115
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 107
: LENGTH: 361
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-995-494-107

Query Match
Best Local Similarity 43.1%; Score 734.5; DB 11; Length 361;
Matches 185; Conservative 24; Mismatches 93; Indels 127; Gaps 12;

QY 6 LLPRLECGAISAHRNLRLPGSSDSPASASPVAGITGM-----CTHARL----- 49
DB 1 LLPRLECGSTIMAHCRKLKLGSDLPASASRYGTTGMKQPTMGCTCAHPCIGLFPVAN 60
QY 50 -----ILYFLVEMFLVAGAGLEPLTSDPSVASASARYRGHARLCLIA 97
DB 61 FERSFYIIGTSPLLDLMLIMTF-----FALLIHVLF- 94
QY 98 NEFCNRYSLMCP--SMSEPLKOSTCLSLPKCMWRRAVPGFLIFELRHRC----- 148
DB 95 -----NRSLCCRGASLFLGLFAFACLSLPRKMDYTGEL-LPLDIFLOVCCLSFYELN 148
QY 149 PTLTDEVQWCDHSSLQDSTPEIKHPASASQVAGTKDMHNTWLIFFIFENELRQSLNS 208
DB 149 TVFOHAEV-----LIFQWMPYLLRWSLHS 172
QY 209 VTQAGVQWRNIGSLQPLPPGFKLPSCLSLSSMDYRR--PRLAN--FVFLVEMGFTWMA 265
DB 173 VAQAGVQWNCIGSLQPLPRFRFRFSCSLSSMDHRHAPCLANLFFFLVDQSFMTLA 232
QY 266 RLILISGPC-DLPASASQASAGITGVSH-----HARLIENFCLFEMES 306
DB 233 RLVNLSAPSGDLPA PASQASAGITGVRHCHWPKSFKFAADSHIGLAFHFAFFFEFFFAVAS 292
QY 307 HSYTQAGVQWPNLGLQPLPPGKRFSCLSLPSMDYGHLPHPANFCIFITRGVSPYLS 366
DB 293 HPIAAGVQWMDLGLSLQPPPGFKQFLCLSLPSMDYRRAPRQANFCIFSRDGVSPCWT 352
QY 367 GWSOTPDLR 375
DB 353 GWSOTPDLR 361

RESULT 12
US-10-007-280A-140
: Sequence 140, Application US/10007280A
: Publication No. US20030059784A1
: GENERAL INFORMATION:
: APPLICANT: Sun, Yongming
: APPLICANT: Recipon, Herve
: APPLICANT: Salceda, Susana
: APPLICANT: Chenghua, Liu
: TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
: FILE REFERENCE: DEX-0257
: CURRENT APPLICATION NUMBER: US/10/007,280A
: CURRENT FILING DATE: 2001-11-07
: PRIOR APPLICATION NUMBER: US 60/246,640
: PRIOR FILING DATE: 2000-11-08
: NUMBER OF SEQ ID NOS: 238
```

```

: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 140
: LENGTH: 449
: TYPE: PRT
: ORGANISM: Homo sapien
US-10-007-280A-140

Query Match
Best Local Similarity 42.6%; Score 729; DB 15; Length 449;
Matches 196; Conservative 40; Mismatches 122; Indels 102; Gaps 19;

QY 6 LLPRLECGAISAHRNLRLPGSSDSPASASPVAGITGMCTHARL-----ILYFLVE 57
DB 1 LLPRLEYGCTILAVCNLHLLPGSSNPTPSASQVAGRDVCHHTWLVCVCVCVCVCVE 60
QY 58 MEFLVAGAGLEPLTSDPSVASASARYRGHARLC-----LANFC--- 100
DB 61 MRFHYVQAGLELLSSDPPISASQAGITGISH---CTWPMHDSFISPGAEPLTFAYTW 117
QY 101 -GRNRSV-----MCPMSPELKOSTCLSLPKMDYRR-----AANPGL---F 139
DB 118 FGRKRIPLTILLIPGPDVLYAFRTELYASPSRQOPASDTARESMGNCNAVPPFLKRW 177
QY 140 ILFLRHRCPTLTODEV-----QWCD--HSSLQDSTPEIKHPA-----SASQVAGTKD 186
DB 178 LIF-----CPFSNOSHMTTKSKWAEVPHPRRAELPAMKQKAKANENSGSVTEPSSSAS 232
QY 187 MHNTWLIFFIFLN-----ELRQSLNSYQAGVQW---RNLGSLQPLPPGFKLPSCLLS 239
DB 233 ILHARMDVY-FLINALIYFLRQSLRSVAQAGVQWCGADLSLQPLPPGFAFPCLSLLS 291
QY 240 SMDYRR--PRLANFEVFLVEMGFTWFARLILISGPCDLPASASQASAGITGVSHAR---- 294
DB 292 SMDYRSLPPCANFANFVFLIETGFHHISQ--ISISAPCDPPASASQASAGITGSHCAQDPVY 350
QY 295 -----LINFCLFEMESHVYTAGVQWPNLGL-----SLQPLPPGLKRSC 334
DB 351 YVSGYIGKODRCYLFEEFFFEETESRTVAAG--RLERSGAISTRSLQPLPPGLKRSC 409
QY 335 LSLPSMDYGHLPHPANFCIFITRGVSPYLSGWSOTPDLR 374
DB 410 LSLSSMDYRCTPPRLAHFCTFSRDGVSPCWSGMSLFDL 449

RESULT 13
US-09-989-920-245
: Sequence 245, Application US/09989920
: Patent No. US20020172957A1
: GENERAL INFORMATION:
: APPLICANT: Macina, Roberto
: APPLICANT: Recipon, Herve
: APPLICANT: Chen, Sei-Yu
: APPLICANT: Sun, Yongming
: APPLICANT: Liu, Chenghua
: TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P
: FILE REFERENCE: DEX-0291
: CURRENT APPLICATION NUMBER: US/09/989,920
: CURRENT FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: 60/252,500
: PRIOR FILING DATE: 2000-11-22
: NUMBER OF SEQ ID NOS: 284
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 245
: LENGTH: 183
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-989-920-245

Query Match
Best Local Similarity 21.1%; Score 428.5; DB 11; Length 183;
Matches 102; Conservative 13; Mismatches 51; Indels 13; Gaps 7;

QY 209 VTQAGVQWRNIGSLQPLPPGFR-LFS--CPSLSSMDYRRPRLANF--FVFLVEMGFT- 262
```


THIS PAGE BLANK (USPTO)


```

OY 181 VAGTKDMHHTWLIETFIENFLRQSLNSTOAGVQWRNIGSLQPLPGKFLSCPSLSS 240
DB 181 VAGTKDMHHTWLIETFIENFLRQSLNSTOAGVQWRNIGSLQPLPGKFLSCPSLSS 240
OY 241 WDYRRPRLANFEVLVENGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
DB 241 WDYRRPRLANFEVLVENGFTMFARLLISGPCDLPASASOSAGITGVSHHARLIFNFC 300
OY 301 LFEEMESHSTOAGVQWRNIGSLQPLPGKFLSCPSLSSMDYGHLPHPANFCIFIRNG 360
DB 301 LFEEMESHSTOAGVQWRNIGSLQPLPGKFLSCPSLSSMDYGHLPHPANFCIFIRNG 360
OY 361 VSPYLSGMSQTPDLR 375
DB 361 VSPYLSGMSQTPDLR 375

RESULT 2
O9NX85 PRELIMINARY: PRT: 152 AA.
ID 09NX85
AC 09NX85:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KAI0536 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsuka T., Sugano S.,
RA Nakamura Y., Isonaga T., Tanaka T.,
RT "NEO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000385; BAA91131.1;
SQ SEQUENCE 152 AA; 16568 MW; 59065F45AANA301B5 CRC64;

Query Match
Best Local Similarity 51.3%; Score 316; DB 4; Length 152;
Matches 80; Conservative 12; Mismatches 38; Indels 26; Gaps 6;

OY 173 HPP-----ASASQVAGTKDMHHTWLIETFIENFLRQSLNSTOAGVQWRNIGSLQPLPG 228
DB 3 HSPRLVANPQV-----LFEISLFEFF-FLNQSALVAQAGVQWRNIGSLQPLPG 53
OY 229 FKLEFCPSLSSMDYRRPRLANFEVLVENGFTMFARL---ILISGPCDLPASASOSAG 285
DB 54 FKQFCSLSSMDYRRPRLANFEVLVENGFTMFARL---ILISGPCDLPASASOSAG 111
OY 286 ITGVSHHAR-----LIFNFCLEEMESHSTOAG 313
DB 112 ITGSHHARQPTSSNPYGIIVFFLPVKTFSGMSQAG 147

RESULT 3
O9H387 PRELIMINARY: PRT: 118 AA.
ID 09H387
AC 09H387:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PRO2550.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,

```

```

RA Xu W., Gao F., Liu M., He F.;
RT "Functional prediction of the coding sequences of 75 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF130089; AAG35515.1;
SQ SEQUENCE 118 AA; 13257 MW; 94688870CAC8760D CRC64;

Query Match
Best Local Similarity 13.4%; Score 273.5; DB 4; Length 118;
Matches 65; Conservative 4; Mismatches 27; Indels 7; Gaps 4;

OY 196 IFENFLRQSLNSTOAGVQWRNIGSLQPLPGKFLSCPSLSSMDYRR-PPRLANFEV 254
DB 4 IFENFLRQSLNSTOAGVQWRNIGSLQPLPGKFLSCPSLSSMDYRR-PPRLANFEV 62
OY 255 FLVENGFTMFARL---ILISGPCDLPASASOSAGITGVSHHAR 294
DB 63 FLVETGFLRVGQAGLELLTSG--DPPASASOSAGITGVSHHTQ 103

RESULT 4
O96AN9 PRELIMINARY: PRT: 73 AA.
ID 096AN9
AC 096AN9:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Similar to hypothetical protein PRO1722.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016905; AAH16905.1;
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8348 MW; 673F61E85300AC2F CRC64;

Query Match
Best Local Similarity 13.0%; Score 263.5; DB 4; Length 73;
Matches 51; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

OY 304 MESHSYQAGVQWRNIGSLQPLPGKFLSCPSLSSMDYGHLPHPAN-FCIFIRGVS 362
DB 1 MESHSYQAGVQWRNIGSLQPLPGKFLSCPSLSSMDYGHLPHPAN-FCIFIRGVS 60
OY 363 PYLSGMSQTPDLR 375
DB 61 PCWPGMSQTPDLR 73

RESULT 5
O9P195 PRELIMINARY: PRT: 118 AA.
ID 09P195
AC 09P195:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PRO1722.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
RA He F.;
RT "Functional prediction of the coding sequences of 79 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";

```

```

RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF119851; AAF69605.1; -
SO SEQUENCE 118 AA; 13067 MW; D6116a242c773055 CRC64;

Query Match
Best Local Similarity 59.2%; Score 262; DB 4; Length 118;
Matches 58; Conservative 6; Pred. No. 6e-19; Indels 6; Gaps 3

QY 208 SVTAQAGVQWRNLGSLQPLPPGFKFLSCGSLSSMDYRR-PPPLANFVFLVMGCTMAR 266
Db 5 SVAQTGVNHNLSLSQPLPPPKQKQSCSLRSSWYTHLPCLANFVFLVETAFRHWQ 64
QY 267 L---ILISGPCDLLPASASQASAGITGVSHHARLIFNFC 301
Db 65 AGKLKLTSG--DQPSASQASAGITISIRTOVGGRFLI 100

RESULT 6
Q96HL9 PRELIMINARY; PRT; 66 AA.
AC Q96HL9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Similar to hypothetical protein PRO1722.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Hmo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLADDER;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
SO SEQUENCE 66 AA; 7409 MW; 77F6F5AFAF7221B3 CRC64;

Query Match
Best Local Similarity 74.2%; Score 261; DB 4; Length 66;
Matches 49; Conservative 4; MisMatches 13; Indels 0; Gaps 0

QY 304 MESHSVTAQAGVQWRNLGSLQPLPPKFKRSCSLSSMDYGHLPHPANCFIFINGVSP 363
Db 1 MESHSVTAQAGVQWRDLGSLQPLPPKQKPSHLSPSSMDYRRVPLIGNFCIFSGGVSP 60
QY 364 YLGGWS 369
Db 61 CWPQWS 66

RESULT 7
Q9HSR3 PRELIMINARY; PRT; 232 AA.
AC Q9HSR3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CDNA: FLJ23147 f1s, clone LNC9295.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Oka T., Suzuki Y.,
RA Obayashi M., Nishii T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK026800; BAB15557.1; -

```

SO	SEQUENCE	232 AA;	26703 MW;	56946BE74C226FF5 CRC64;
	Query Match	12.7%;	Score 259;	DB 4; Length 232;
	Best Local Similarity	67.6%;	Pred. No. 2.7e-18;	
	Matches	50; Conservative	5; Mismatches 19;	Indels 0; Gaps 0;
QY	302	FEMESHVSTQGVQWVWPNLGSLOPLPPGKLRFSCLSPSSMDYR-RPRLANFVFLVEMGFTMFAR	361	
	:::	:::	:::	
Db	159	WQTSRSYVQGVQWQCDCLLOPPRPKRFRSCSLSPSSMDIRRPRLANFCIISRGV	218	
QY	362	SPYLGSWSQTPDLR	375	
Db	219	SPCPWGSMSRCTDLR	232	
RESULT 8				
Q9H5D5	ID	Q9H5D5	PRELIMINARY;	PRT; 162 AA.
AC	Q9H5D5;			
DT	01-MAR-2001 (TREMBLrel. 16; Created)			
DT	01-MAR-2001 (TREMBLrel. 16; Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19; Last annotation update)			
DE	CDNA: FLJ23555 fis, clone LNC09438.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RA	SEQUENCE FROM N.A.			
RP	TISSUE=LUNG;			
RC	Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,			
RA	Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,			
RA	Yanada K., Fujii Y., Ozaki K., Hiraio M., Ohmori Y., Ota T., Suzuki Y.,			
RA	Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,			
RA	Isgogal T., Sugano S.,			
RT	"NEO human cDNA sequencing project."			
RL	Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AK027208; BAB15692.1; "			
SO	SEQUENCE	162 AA; 17632 MW; C9A857907E44D2E8 CRC64;		
	Query Match	12.6%;	Score 257;	DB 4; Length 162;
	Best Local Similarity	47.0%;	Pred. No. 2.8e-18;	
	Matches	63; Conservative	2; Mismatches 23;	Indels 46; Gaps 2;
QY	208	SVTQAGVQWVWNLGSLQPLPPGKFLKSCPSLSSMDYR-RPRLANFVFLVEMGFTMFAR	266	
	:			
Db	63	SVPRAGVQWCMNGSLQPLPPGKFLKSCPSLSSMDYR-RPRLANFVFLVEMGFTMFAR	111	
QY	267	LILISPCDLPASASQSGAIGVSHNHLINFLCFEMESHVSTQGVQWVWPNLGSLOPL	326	
Db	112	FFEMESGVPHAGVQWMDLGSQOPR	137	
QY	327	PGLKRFSCSLSPSS	340	
Db	138	PGFKRFSYLSLSPSS	151	
RESULT 9				
Q8TB48	ID	Q8TB48	PRELIMINARY;	PRT; 72 AA.
AC	Q8TB48;			
DT	01-JUN-2002 (TREMBLrel. 21; Created)			
DT	01-JUN-2002 (TREMBLrel. 21; Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21; Last annotation update)			
DE	Similar to PR02550.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RA	SEQUENCE FROM N.A.			
RP	TISSUE=TESTIS;			
RA	Strausberg R.;			

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021822; AH21822.1; -
DR InterPro: IPR000822; Znf.C2H2.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
KM Hypothetical protein.
SQ SEQUENCE 238 AA; 27158 MW; FACPDDB5BEICIF CRC64;
Query Match 12.3%; Score 250; DB 4; Length 238;
Best Local Similarity 45.5%; Pred. No. 2,3e-17;
Matches 60; Conservative 7; Mismatches 23; Indels 42; Gaps 3;
QY 1 MEPSLLI-PRLENGAISAHRLRLPGSSDPSASAPVAGITGCTHARLILYFLVEME 59
Db 96 LKWSLTSPKLECGAIVCHNRLCGSSDLSAQAGIAGACHAQLI-FVELVETG 154
QY 60 FLHVAGAGLELPISDDPSVASOSARIRGTHARLILANCGRRVSLMCPSPSELKOS 119
Db 155 FHHFDAGFELLTSSDPALASQA----- 179
QY 120 TCULSPKMDYR 131
Db 180 -----PKCMDYK 186

RESULT 14
Q9HA67 PRELIMINARY; PRT; 151 AA.

AC Q9HA67:
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CDNA FLJ12155 fis, clone MAMMA1000472.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY GLAND;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagaatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Minomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK022217; BAB13989.1; -
SQ SEQUENCE 151 AA; 16972 MW; ECE55BBD50DF2811 CRC64;

Query Match 12.2%; Score 248; DB 4; Length 151;
Best Local Similarity 40.1%; Pred. No. 2,2e-17;
Matches 75; Conservative 10; Mismatches 52; Indels 50; Gaps 9;

QY 146 HRCPTLTQDEVQCDHSSLOPSTPEIKHPASASQVAGTMDMHHYTMILIFIFENFLROS 205
Db 10 HRC-----ICKHYA-PTTAPHL-----FFETE 30
QY 206 LNSVTQAGVQWRNIGSLQPLPGFKLSCPSLLSMDYRRPR-LANFEVFLVEMGFTMF 264
Db 31 SHSVTQAGVQWCDLGSLOPSPGFKQFSCSLSRMDYRRVPLCLANFIVELVETGRCRV 90
QY 265 --ARLLISGRCULPASASASAGITGVSH-ARLIFNCLFEMESHVYQAGVQWPNLGS 321
Db 91 GQGLKLLTS-SDLPASACOSAG--DYRHEPLRLALTLCHF-----ISRTCTSDVFYICRD 143

QY 322 LQPLPG 328
Db 144 LERIPHG 150

RESULT 15
Q9HBS7 PRELIMINARY; PRT; 130 AA.

ID Q9HBS7:
AC Q9HBS7:
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Hypothetical 14.2 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu J.R., Man D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel Human cDNA clones with function of inhibiting cancer cell
growth."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF217973; AAC17216.1; -
KM Hypothetical protein.
SQ SEQUENCE 130 AA; 14153 MW; FF338649055F70FA CRC64;

Query Match 12.2%; Score 247.5; DB 4; Length 130;
Best Local Similarity 59.2%; Pred. No. 2e-17;
Matches 58; Conservative 5; Mismatches 34; Indels 1; Gaps 1;

QY 2 EFSLILPRLECGAIVCHNRLCGSSDPSASAPVAGITGCTHARLILYFLVEMEFL 61
Db 33 EVSLSPRLCSGTSIAHCHNCLGLSSYSVASASRYTGIGMCHHAQLI-FVELVMEFR 91
QY 62 HVGAGAGLELPISDDPSVASOSARIRGTHARLILANF 99
Db 92 HVGOTSELLASSSPASASASAGITGVSHCAMPLAAY 129

Search completed: July 17, 2003, 18:19:21
Job time : 83 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 17, 2003, 18:07:48 ; Search time 23 Seconds
(without alignments)
676.244 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MEFSLPRLCNCALSAHR.....FIRGVSPLYSGWSTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112392

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	595	29.3	587	1	ALU2_HUMAN
2	568	27.9	587	1	ALU3_HUMAN
3	549.5	27.0	591	1	ALU8_HUMAN
4	498.5	24.5	591	1	ALU1_HUMAN
5	389	19.1	585	1	ALU5_HUMAN
6	357.5	17.6	593	1	ALU6_HUMAN
7	356.5	17.5	603	1	ALU4_HUMAN
8	330.5	16.2	593	1	ALU7_HUMAN
9	161	7.9	841	1	NEK4_HUMAN
10	101.5	5.0	438	1	SGCE_HUMAN
11	101	5.0	717	1	MTOL_HUMAN
12	100	4.9	629	1	2195_HUMAN
13	98.5	4.8	741	1	RED1_HUMAN
14	98	4.8	501	1	GGG2_HUMAN
15	96.5	4.7	1643	1	HNPA_HUMAN
16	91.5	4.5	631	1	HNPA_HUMAN
17	91.5	4.5	1776	1	POLR_OYAV
18	90.5	4.4	282	1	VMSA_MHYV6
19	90	4.4	431	1	VMSA_MHYV8
20	88	4.3	209	1	HS27_CANFA
21	88	4.3	726	1	ADDB_HUMAN
22	87	4.3	1153	1	NS2A_HUMAN
23	86.5	4.3	961	1	BASO_MOUSE
24	86.5	4.3	1886	1	GP21_RAT
25	85.5	4.2	832	1	DPOL_HPBVA
26	84	4.1	366	1	VMSA_HPBVA
27	84	4.1	843	1	DPOL_HPBVA
28	83.5	4.1	1304	1	S3B1_MOUSE
29	83	4.1	716	1	HGFL_MOUSE
30	83	4.1	763	1	DPOL_HPBVA
31	83	4.1	1216	1	ABCP_RAT
32	82.5	4.1	481	1	DPOL_HPBVA
33	82.5	4.1	845	1	DPOL_HPBVA

34	82.5	4.1	1304	1	S3B1_HUMAN	075533 homo sapien
35	82	4.0	226	1	VMSA_HPBVA	P31869 hepatitis b
36	82	4.0	832	1	DPOL_HPBVA	P03156 hepatitis b
37	81.5	4.0	389	1	VMSA_HPBVA	P03138 hepatitis b
38	81.5	4.0	597	1	C4BP_HUMAN	P04003 homo sapien
39	81.5	4.0	843	1	DPOL_HPBVA	P17393 hepatitis b
40	81	4.0	723	1	TERM_ADECR	O96682 canine aden
41	81	4.0	901	1	Y298_HUMAN	O15016 homo sapien
42	80.5	4.0	389	1	VMSA_HPBVA	P12911 hepatitis b
43	80.5	4.0	413	1	GAT1_HUMAN	P15976 homo sapien
44	80.5	4.0	1299	1	JAK3_MOUSE	O62137 mus musculu
45	79.5	3.9	619	1	REL_HUMAN	O04864 homo sapien

ALIGNMENTS

RESULT 1
ID ALU2_HUMAN STANDARD: PRT: 587 AA.
AC P39189:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alu subfamily SB sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=95021758; PubMed=7935834;
RA Claverie J.-M., Makalowski W.;
RT "Alu alert."
RL Nature 371:752-752(1994).
RN [2]
RP CONCEPT.
RP MEDLINE=92241891; PubMed=1572661;
RA Claverie J.-M.;
RT "Identifying coding exons by similarity search: alu-derived and other
potentially misleading protein sequences.";
RL Genomics 12:838-841(1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=88333009; PubMed=3138422;
RA "The Alu family developed through successive waves of fixation
Quentin Y."
RT closely connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202(1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=91178815; PubMed=1706781;
RA Jurka J., Miosavljetic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121(1991).
CC -I- MISCELLANEOUS: VARIOUS ANALYSES (SEE REF.3 AND REF.4) INDICATE
THAT ALU REPEATS FALL INTO 8 SUBFAMILIES. THEREFORE, 8 ALU WARNING
FRAMES CONCEPTUAL TRANSLATIONS OF EACH OF THESE CLASSES OF ALU
REPEATS.
CC -I- MISCELLANEOUS: ISOLATED 'X' INDICATES THE PRESENCE OF A STOP
CODON, 'XXX' IS USED TO SEPARATE THE VARIOUS TRANSLATION PHASES.
CC -I- CAUTION: THIS ALU ENTRY IS PROVIDED IN ORDER TO AVOID THE FURTHER
POLLUTION OF PROTEIN SEQUENCE DATABASES WITH ALU-DERIVED AMINO
ACID SEQUENCES.
CC -I- CAUTION: ALU REPEATITIVE SEQUENCES ARE INTERSPERSED IN HUMAN AND
PRIMATE GENOMES WITH AN AVERAGE SPACING OF 4 KB. SOME OF THEM ARE
ACTIVELY TRANSCRIBED BY POL III. NORMAL TRANSCRIPTS MAY CONTAIN
ALU-DERIVED SEQUENCES IN 5' OR 3' UNTRANSLATED REGIONS. HOWEVER,
CDNA LIBRARIES ALSO CONTAIN PARTIAL AND/OR REARRANGED CDNAS
LIGATED WITH ALU-DERIVED SEQUENCE IN ANY ORIENTATION. ALTHOUGH ALU
ELEMENTS (ESPECIALLY SITUATED ON THE COMPLEMENTARY STRAND) HAVE A
GREAT POTENTIAL TO CREATE ADDITIONAL/ALTERNATIVE EXONS,

CC CONSIDERATION SHOULD BE GIVEN TO THE POSSIBILITY THAT THE PRESENCE
 CC OF AN ALU IN AN OPEN READING FRAME MAY HAVE RESULTED FROM A
 CC CLONING ARTIFACT OR MAY BE DUE TO MISINTERPRETATION OF SEQUENCING
 CC DATA. THIS POINT HAS BEEN OVERLOOKED ON SEVERAL OCCASIONS, WITH
 CC THE CONSEQUENCE OF ERRONEOUS ALU-DERIVED AMINO ACID SEQUENCES
 CC BEING REPORTED.
 CC -1- CAUTION: ANY SIGNIFICANT SIMILARITY OF A PUTATIVE PROTEIN SEQUENCE
 CC WITH AN ALU-TRANSLATED ENTRY MUST BE TAKEN AS A WARNING THAT A
 CC PART OF ALU REPEAT MAY HAVE BEEN ARTIFACTUALLY INCLUDED IN THE
 CC CODING NUCLEOTIDE SEQUENCE.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U14568; -; NOT_ANNOTATED_CDS.
 CC
 CC Hypothetical protein.
 CC FT DOMAIN 1 96 FRAME-1.
 CC FT DOMAIN 100 194 FRAME-2.
 CC FT DOMAIN 198 292 FRAME-3.
 CC FT DOMAIN 296 391 FRAME-4.
 CC FT DOMAIN 395 489 FRAME-5.
 CC FT DOMAIN 493 587 FRAME-6.
 CC SEQUENCE 587 AA; 63703 MW; 3EAB3EE33929203 CRC64;
 CC
 CC Query Match 29.3%; Score 595; DB 1; Length 587;
 CC Best Local Similarity 41.5%; Pred. No. 1,1e-45;
 CC Matches 153; Conservative 14; Mismatches 96; Indels 106; Gaps 8;

QY 6 LLLPLECGAISARNRLRPPSSDPASASPVAGITGCHTARLLTFVLEMEFLHVGQ 65
 DB 304 LSPRLCSGSAISACKRLRPGSRHSPASASRVAGTGRHARLI-FVFLVETGFHRYSQ 362
 QY 66 AGLEPFSDDPSVASOSARRYRTGTHARLCLANFCGRNRVLMCPSPMSPELKSTCLSP 125
 DB 363 DGLDLT-----SASAR-----LGLP 378
 QY 126 KCMYRRAAVPGLFLEFLRHRCPTLQDEVQWCDHSSLPSTPEIKHPASASQVAGTK 185
 DB 379 KCMYRRPRPRPAXXFP-----XDGVSIC-----RPGSAVARSRLTAGSAS--- 421
 QY 186 DMHYTWLIFIFINFLRQSLNSTVQAGVQWRNIGSLQRPGRKLPSCPLSSWDRR 245
 DB 422 -----RVHAILPPEXGLQAPATTP----- 444
 QY 246 PPRLANFEVLVENGFTMFARLLISGPCDLPASASAGITGVSHARLIFNCLPEME 305
 DB 445 ---GXFLYFXRRGFVLAHWVISXRPDPASASQAGITGVSHARXXX---FEETE 497
 QY 306 SHSTVQAGVQWRNIGSLQRPGLKRSCLSPSSWDYGHLPHPANFCITRGVSPYL 365
 DB 498 SRSVAQAGVQWRDGLQAPPPGFTPFSCSLSPSSWDYRRPPRPANFCIFSRDGVSPXC 557
 QY 366 SGWSQTPDL 374
 DB 558 PGWSRSPDL 566

RESULT 2
 ALU3_HUMAN
 ID ALU3_HUMAN STANDARD: PRT: 587 AA.
 AC P39190:
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alu subfamily SBI sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021758; PubMed=7935834;
 RA Claverie J.-M., Makalowski W.;
 RT "Alu alert";
 RL Nature 371:752-752(1994).
 RN [2]
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 RT potentially misleading protein sequences.";
 RL Genomics 12:838-841(1992).
 RN [3]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation
 RT closely connected with primate lineage history.";
 RL J. Mol. Evol. 27:194-202(1988).
 RN [4]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Miosavljetic A.;
 RT "Reconstruction and analysis of human Alu genes.";
 RL J. Mol. Evol. 32:105-121(1991).
 CC -1- MISCELLANEOUS: VARIOUS ANALYSES (SEE REF.3 AND REF.4) INDICATE
 CC THAT ALU REPEATS FALL INTO 8 SUBFAMILIES. THEREFORE, 8 ALU WARNING
 CC CONSENSUS SEQUENCES HAVE BEEN CONSTITUTED THAT CONTAIN ALL SIX
 CC FRAMES CONCEPTUAL TRANSLATIONS OF EACH OF THESE CLASSES OF ALU
 CC REPEATS.
 CC -1- MISCELLANEOUS: ISOLATED 'X' INDICATES THE PRESENCE OF A STOP
 CC CODON. 'XXX' IS USED TO SEPARATE THE VARIOUS TRANSLATION PHASES.
 CC -1- CAUTION: THIS ALU ENTRY IS PROVIDED IN ORDER TO AVOID THE FURTHER
 CC POLLUTION OF PROTEIN SEQUENCE DATABASES WITH ALU-DERIVED AMINO
 CC ACID SEQUENCES.
 CC -1- CAUTION: ALU REPEAT SEQUENCES ARE INTERSPERSED IN HUMAN AND
 CC PRIMATE GENOMES WITH AN AVERAGE SPACING OF 4 KB. SOME OF THEM ARE
 CC ACTIVELY TRANSCRIBED BY POL. III. NORMAL TRANSCRIPTS MAY CONTAIN
 CC ALU-DERIVED SEQUENCES IN 5' OR 3' UNTRANSLATED REGIONS. HOWEVER,
 CC CDNA LIBRARIES ALSO CONTAIN PARTIAL AND/OR REARRANGED CDNAS
 CC LIGATED WITH ALU-DERIVED SEQUENCE IN ANY ORIENTATION. ALTHOUGH ALU
 CC ELEMENTS (ESPECIALLY SITUATED ON THE COMPLEMENTARY STRAND) HAVE A
 CC GREAT POTENTIAL TO CREATE ADDITIONAL/ALTERNATIVE EXONS,
 CC CONSIDERATION SHOULD BE GIVEN TO THE POSSIBILITY THAT THE PRESENCE
 CC OF AN ALU IN AN OPEN READING FRAME MAY HAVE RESULTED FROM A
 CC CLONING ARTIFACT OR MAY BE DUE TO MISINTERPRETATION OF SEQUENCING
 CC DATA. THIS POINT HAS BEEN OVERLOOKED ON SEVERAL OCCASIONS, WITH
 CC THE CONSEQUENCE OF ERRONEOUS ALU-DERIVED AMINO ACID SEQUENCES
 CC BEING REPORTED.
 CC -1- CAUTION: ANY SIGNIFICANT SIMILARITY OF A PUTATIVE PROTEIN SEQUENCE
 CC WITH AN ALU-TRANSLATED ENTRY MUST BE TAKEN AS A WARNING THAT A
 CC PART OF ALU REPEAT MAY HAVE BEEN ARTIFACTUALLY INCLUDED IN THE
 CC CODING NUCLEOTIDE SEQUENCE.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U14569; -; NOT_ANNOTATED_CDS.
 CC
 CC Hypothetical protein.
 CC FT DOMAIN 1 96 FRAME-1.
 CC FT DOMAIN 100 194 FRAME-2.
 CC FT DOMAIN 198 292 FRAME-3.
 CC FT DOMAIN 296 391 FRAME-4.
 CC FT DOMAIN 395 489 FRAME-5.
 CC FT DOMAIN 493 587 FRAME-6.
 CC SEQUENCE 587 AA; 63573 MW; 85C4155726DEF235 CRC64;

Query Match 27.9%; Score 568; DB 1; Length 587;
 Best Local Similarity 41.5%; Pred. No. 2.9e-43;
 Matches 153; Conservative 15; Mismatches 95; Indels 106; Gaps 10;

OY 6 LLPRLECGAISAHNRLRPGSSDSPASAPVAGITGCTHARLLLYFLVMEFLHVGQ 65
 DB 304 LSPRLCSCGSAHCKRLPGSRHSPASASQVAGTGAHRYARLL-FVFLVETGHHRRSR 362
 OY 66 AGLELPTSDSDPSASARRTGTHARLLCLANFCGRNRYSLMCSHSPDLKOSTCLSLP 125
 DB 363 DGLDLT-----SXSA-----LGLP 378
 OY 126 KCMVRAAAGLFLFRLHRCPTLTODEVOMCHSLQSTPEIKHPASASQVAGTK 185
 DB 379 KCMVRRPRPRAXXFF-----XDGVSIC-----RPGMSA--VAASR 414
 OY 186 DMHHTWLIFFIFNFIROSLSNVTQAGVOMRNLGSLQPLPPGFKLSLSPSSMDYRR 245
 DB 415 -----LTASSASRVHA-----ILPPPKXKLGIAQAPATP----- 444
 OY 246 PPRLANFVFLVEMGTFMFAILLISGPCDLPASASQAGITGVSHHARLLFNCLFEME 305
 DB 445 -----GXFLYEXXRRGFTVLGAGVSIKPRDPASASQAGITGVSHHARXXX--FFETE 497
 OY 306 SHSVTQAGVOMRNLGSLQPLPPGLKRFSCSLSPSSMDYGHLPHPANFCIFRGGVSPYL 365
 DB 498 SRSVAQAGVOMRDLGSLQAPPPGFTPFSCSLSPSSMDYRRPPLRANCFISRDGVSFFX 557
 OY 366 SGMSQTPDL 374
 DB 558 PGMSRSPDL 566

RESULT 3

ALU8_HUMAN STANDARD; PRT; 591 AA.
 ID AC P39195;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alu subfamily SX sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=95021758; PubMed=7935834;
 RA Claverie J.-M., Makalowski W.,
 RT "Alu alert."
 RL Nature 371:752-752(1994).
 RN [2]
 RP CONCEPT.
 RP MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 RT potentially misleading protein sequences."
 RL Genomics 12:838-841(1992).
 RN [3]
 RP ALU FAMILIES CLASSIFICATION.
 RP MEDLINE=86333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation
 RT closely connected with primate lineage history."
 RL J. Mol. Evol. 27:194-202(1988).
 RN [4]
 RP ALU FAMILIES CLASSIFICATION.
 RP MEDLINE=91178815; PubMed=1106781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes."
 RL J. Mol. Evol. 32:105-121(1991).
 CC -I- MISCELLANEOUS; VARIOUS ANALYSES (SEE REF.3 AND REF.4) INDICATE

CC THAT ALU REPEATS FALL INTO 8 SUBFAMILIES. THEREFORE, 8 ALU WARNING
 CC CONSENSUS SEQUENCES HAVE BEEN CONSTITUTED THAT CONTAIN ALL SIX
 CC FRAMES CONCEPTUAL TRANSLATIONS OF EACH OF THESE CLASSES OF ALU
 CC REPEATS.
 CC -I- MISCELLANEOUS; ISOLATED 'X' INDICATES THE PRESENCE OF A STOP
 CC CODON, 'XXX' IS USED TO SEPARATE THE VARIOUS TRANSLATION PHASES.
 CC -I- CAUTION: THIS ALU ENTRY IS PROVIDED IN ORDER TO AVOID THE FURTHER
 CC POLLUTION OF PROTEIN SEQUENCE DATABASES WITH ALU-DERIVED AMINO
 CC ACID SEQUENCES.
 CC -I- CAUTION: ALU REPETITIVE SEQUENCES ARE INTERSPERSED IN HUMAN AND
 CC PRIMATE GENOMES WITH AN AVERAGE SPACING OF 4 KB. SOME OF THEM ARE
 CC ACTIVELY TRANSCRIBED BY POL III. NORMAL TRANSCRIPTS MAY CONTAIN
 CC ALU-DERIVED SEQUENCES IN 5' OR 3' UNTRANSLATED REGIONS. HOWEVER,
 CC CDNA LIBRARIES ALSO CONTAIN PARTIAL AND/OR REARRANGED CDNAS
 CC LOCATED WITH ALU-DERIVED SEQUENCE IN AN ORIENTATION, ALTHOUGH ALU
 CC ELEMENTS (ESPECIALLY SITUATED ON THE COMPLEMENTARY STRAND) HAVE A
 CC GREAT POTENTIAL TO CREATE ADDITIONAL/ALTERNATIVE EXONS,
 CC CONSIDERATION SHOULD BE GIVEN TO THE POSSIBILITY THAT THE PRESENCE
 CC OF AN ALU IN AN OPEN READING FRAME MAY HAVE RESULTED FROM A
 CC CLONING ARTIFACT OR MAY BE DUE TO MISINTERPRETATION OF SEQUENCING
 CC DATA. THIS POINT HAS BEEN OVERLOOKED ON SEVERAL OCCASIONS, WITH
 CC THE CONSEQUENCE OF ERRONEOUS ALU-DERIVED AMINO ACID SEQUENCES
 CC BEING REPORTED.
 CC -I- CAUTION: ANY SIGNIFICANT SIMILARITY OF A PUTATIVE PROTEIN SEQUENCE
 CC WITH AN ALU-TRANSLATED ENTRY MUST BE TAKEN AS A WARNING THAT A
 CC PART OF ALU REPEAT MAY HAVE BEEN ARTIFACTUALLY INCLUDED IN THE
 CC CODING NUCLEOTIDE SEQUENCE.

CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb-sib.ch/announce/>
 CC OR SEND AN EMAIL TO license@sib.ch).

CC EMBL; U14574; ; NOT_ANNOTATED_CDS.

CC DR EMBL; U14574; ; NOT_ANNOTATED_CDS.
 CC HYPOTHETICAL PROTEIN.
 CC FT DOMAIN 1 96 FRAME-1.
 CC FT DOMAIN 100 195 FRAME-2.
 CC FT DOMAIN 189 284 FRAME-3.
 CC FT DOMAIN 298 393 FRAME-4.
 CC FT DOMAIN 397 492 FRAME-5.
 CC FT DOMAIN 496 591 FRAME-6.
 CC SQ SEQUENCE 591 AA; 64395 MW; AC8154AD8A6BB280 CRC64;

Query Match 27.0%; Score 549.5; DB 1; Length 591;
 Best Local Similarity 38.9%; Pred. No. 1.3e-41;
 Matches 150; Conservative 15; Mismatches 84; Indels 137; Gaps 11;

OY 6 LLPRLECGAISAHNRLRPGSSDSPASAPVAGITGCTHARLLLYFLVMEFLHVGQ 65
 DB 306 LSPRLCSCGSAHCKRLRPGSSDSPASAPVAGITGAHRYARLL-FVFLVETGHHRRSR 364
 OY 66 AGLELPTSDSD--PSYASQASARYRTGTHARLLCLANFCGRNRYSLMCPMS-----P 114
 DB 365 AGLELTSGLDLPKXYLGIAQATARG---LBRXXXFFXDGVSICLRPGMSANVARSRLTAT 420
 OY 115 ELKOSTCLSLPKCMD---YRRAAVPGLFLEFLHRCPTLTODEVOMCHSLQSTPE 170
 DB 421 SASRVOATILPQRPXKLGIAQATATPGXFLYXXHRCFTMLAR-----LVNSNXPQ 471
 OY 171 IKHPPASASQVAGTDMHHTWLIFFIFNFIROSLSNVTQAGVOMRNLGSLQPLPPGPK 230
 DB 472 V-----IF----- 474
 OY 231 LFSCLSSMDYRR-PPRLANFVFLVEMGTFMFAILLISGPCDLPASASQAGITGV 289
 DB 475 -----LPCMDYRRPR-----PASAXXX----- 495
 OY 290 SHHARLLNFCLEFESHSHSVTQAGVOMRNLGSLQPLPPGLKRFSCSLSPSSMDYGHLP 349
 DB 496 -----FFETESRSVAQAGVOMRDLGSLQAPPPGFTPFSCSLSPSSMDYRRAPR 544

RT "Identifying coding exons by similarity search: alu-derived and other
RT potentially misleading protein sequences.";
RT Genomics 12:838-841(1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE-88333009; PubMed-3138422;
RA "Quentin Y.;"
RT "The Alu family developed through successive waves of fixation
RT closely connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202(1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE-91178815; PubMed-1706781;
RA Jurka J., Miosavljjevic A.;"
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121(1991).
CC -I- MISCELLANEOUS: VARIOUS ANALYSES (SEE REF.3 AND REF.4) INDICATE
CC THAT ALU REPEATS FALL INTO 8 SUBFAMILIES. THEREFORE, 8 ALU WARNING
CC CONSENSUS SEQUENCES HAVE BEEN CONSTITUTED THAT CONTAIN ALL SIX
CC FRAMES CONCEPTUAL TRANSLATIONS OF EACH OF THESE CLASSES OF ALU
CC REPEATS.
CC -I- MISCELLANEOUS: ISOLATED 'X' INDICATES THE PRESENCE OF A STOP
CC CODON, 'XXX' IS USED TO SEPARATE THE VARIOUS TRANSLATION PHASES.
CC -I- CAUTION: THIS ALU ENTRY IS PROVIDED IN ORDER TO AVOID THE FURTHER
CC POLLUTION OF PROTEIN SEQUENCE DATABASES WITH ALU-DERIVED AMINO
CC ACID SEQUENCES.
CC -I- CAUTION: ALU REPEAT SEQUENCES ARE INTERSPERSED IN HUMAN AND
CC PRIMATE GENOMES WITH AN AVERAGE SPACING OF 4 KB. SOME OF THEM ARE
CC ACTIVELY TRANSCRIBED BY POL III. NORMAL TRANSCRIPTS MAY CONTAIN
CC ALU-DERIVED SEQUENCES IN 5' OR 3' UNTRANSLATED REGIONS. HOWEVER,
CC CDNA LIBRARIES ALSO CONTAIN PARTIAL AND/OR REARRANGED CDNAS
CC LINKED WITH ALU-DERIVED SEQUENCE IN AN ORIENTATION. ALTHOUGH ALU
CC ELEMENTS (ESPECIALLY SITUATED ON THE COMPLEMENTARY STRAND) HAVE A
CC GREAT POTENTIAL TO CREATE ADDITIONAL/ALTERNATIVE EXONS,
CC CONSIDERATION SHOULD BE GIVEN TO THE POSSIBILITY THAT THE PRESENCE
CC OF AN ALU IN AN OPEN READING FRAME MAY HAVE RESULTED FROM A
CC CLONING ARTIFACT OR MAY BE DUE TO MISINTERPRETATION OF SEQUENCING
CC DATA. THIS POINT HAS BEEN OVERLOOKED ON SEVERAL OCCASIONS, WITH
CC THE CONSEQUENCE OF ERRONEOUS ALU-DERIVED AMINO ACID SEQUENCES
CC BEING REPORTED.
CC -I- CAUTION: ANY SIGNIFICANT SIMILARITY OF A POTENTIAL PROTEIN SEQUENCE
CC WITH AN ALU-TRANSLATED ENTRY MUST BE TAKEN AS A WARNING THAT A
CC PART OF ALU REPEAT MAY HAVE BEEN ARTIFICIALLY INCLUDED IN THE
CC CODING NUCLEOTIDE SEQUENCE.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC
CC EMBL: U14571; -; NOT_ANNOTATED_CDS.
KW Hypothetical protein.
FT DOMAIN 1 95 FRAME-1.
FT DOMAIN 2 99 193 FRAME-2.
FT DOMAIN 3 197 291 FRAME-3.
FT DOMAIN 4 295 389 FRAME-4.
FT DOMAIN 5 393 487 FRAME-5.
FT DOMAIN 6 491 585 FRAME-6.
SQ SEQUENCE 585 AA; 63957 MW; 46EE8C4FA93650A7 CRC64;
Query Match 19.1% Score 389; DB 1; Length 585;
Best Local Similarity 33.9% Pred No. 3e-27;
Matches 130; Conservative 23; Mismatches 100; Indels 130; Gaps 17;
QY 17 SAHNRRLRLPGSSDSDASPVAGITGKTHARLLYFLVLEWFLHVGAGLELPTSDP 76
DB 189 SVSASQXRYRTGTHARLCLANFCR-----NRV-----SLKC-PSWSPELKOST 120
QY 17 SAHNRRLRLPGSSDSDASPVAGITGKTHARLLYFLVLEWFLHVGAGLELPTSDP 76
DB 189 SVSASQXRYRTGTHARLCLANFCR-----NRV-----SLKC-PSWSPELKOST 120

DB 233 SLKTIQKLAGRC--ARL-XSOLLGRLEQENRLNMGCGGCGSEPRSRHCTPAN---ROSE 285
QY 121 CLSLPKCMCDYRAAVPGLFLEFLRHRCPTLTQDEVMCDHSLQPTSTPEIKHPASASQ 180
DB 286 TPSSQK-----XXFFLRSL-----ALSPGW-----SA 309
QY 181 VAGTDMHHYTWLFIFFLNFLOSINSYVAGVQWRNIGSLQPLP---GKFLFSCPSL 237
DB 310 VA-----RSRLTGTASARVQ-----ALLPQPEXELGID----- 338
QY 238 LSSWDYRRPRLANFVLEWGFTHARLLISPCDLPASASAGITGVSHARLIF 297
DB 339 -----ARATTPSXLVFXXRGRFTMLARVMSISXPRPPASASAGITGVSHARLXXX 392
QY 298 NF-----CLFEMESHVTOAGVQWRNIGSLQPLPGLKRFSCLSLPSMDVGHLPHPA 351
DB 393 FFYDGVSLC-----RQAGVQWRNIGSLQPLPGLKRFSCLSLPSMDVGHLPHPA 443
QY 352 NFCEIFRIGVSPYLSGWSQTPDL 374
DB 444 NFCEIFRIGVSPYLSGWSQTPDL 466
RESULT 6
ALU6_HUMAN STANDARD; PRT; 593 AA.
AC P39193;
DT 01-FEB-1995 (rel. 31, Created)
DT 01-FEB-1995 (rel. 31, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Alu subfamily SP sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95021758; PubMed-7935834;
RA Claverie J.-M., Makalowski W.;"
RT "Alu alert.";
RL Nature 371:752-752(1994).
RN [2]
RP CONCEPT.
RX MEDLINE-92241891; PubMed-1572661;
RA Claverie J.-M.;"
RT "Identifying coding exons by similarity search: alu-derived and other
RT potentially misleading protein sequences.";
RL Genomics 12:838-841(1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE-88333009; PubMed-3138422;
RA "Quentin Y.;"
RT "The Alu family developed through successive waves of fixation
RT closely connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202(1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE-91178815; PubMed-1706781;
RA Jurka J., Miosavljjevic A.;"
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121(1991).
CC -I- MISCELLANEOUS: VARIOUS ANALYSES (SEE REF.3 AND REF.4) INDICATE
CC THAT ALU REPEATS FALL INTO 8 SUBFAMILIES. THEREFORE, 8 ALU WARNING
CC CONSENSUS SEQUENCES HAVE BEEN CONSTITUTED THAT CONTAIN ALL SIX
CC FRAMES CONCEPTUAL TRANSLATIONS OF EACH OF THESE CLASSES OF ALU
CC REPEATS.
CC -I- MISCELLANEOUS: ISOLATED 'X' INDICATES THE PRESENCE OF A STOP
CC CODON, 'XXX' IS USED TO SEPARATE THE VARIOUS TRANSLATION PHASES.
CC -I- CAUTION: THIS ALU ENTRY IS PROVIDED IN ORDER TO AVOID THE FURTHER
CC POLLUTION OF PROTEIN SEQUENCE DATABASES WITH ALU-DERIVED AMINO
CC ACID SEQUENCES.
CC -I- CAUTION: ALU REPEAT SEQUENCES ARE INTERSPERSED IN HUMAN AND
CC PRIMATE GENOMES WITH AN AVERAGE SPACING OF 4 KB. SOME OF THEM ARE

CC ACTIVELY TRANSCRIBED BY POL. III. NORMAL TRANSCRIPTS MAY CONTAIN
 CC ALU-DERIVED SEQUENCES IN 5' OR 3' UNTRANSLATED REGIONS. HOWEVER,
 CC CDNA LIBRARIES ALSO CONTAIN PARTIAL AND/OR REARRANGED CNAS
 CC LIGATED WITH ALU-DERIVED SEQUENCE IN ANY ORIENTATION. ALTHOUGH ALU
 CC ELEMENTS (ESPECIALLY SITUATED ON THE COMPLEMENTARY STRAND) HAVE A
 CC GREAT POTENTIAL TO CREATE ADDITIONAL/ALTERNATIVE EXONS,
 CC CONSIDERATION SHOULD BE GIVEN TO THE POSSIBILITY THAT THE PRESENCE
 CC OF AN ALU IN AN OPEN READING FRAME MAY HAVE RESULTED FROM A
 CC CLONING ARTIFACT OR MAY BE DUE TO MISINTERPRETATION OF SEQUENCING
 CC DATA. THIS POINT HAS BEEN OVERLOOKED ON SEVERAL OCCASIONS, WITH
 CC THE CONSEQUENCE OF ERRONEOUS ALU-DERIVED AMINO ACID SEQUENCES
 CC BEING REPORTED.
 CC -1- CAUTION: ANY SIGNIFICANT SIMILARITY OF A PUTATIVE PROTEIN SEQUENCE
 CC WITH AN ALU-TRANSLATED ENTRY MUST BE TAKEN AS A WARNING THAT A
 CC PART OF ALU REPEAT MAY HAVE BEEN ARTIFACTUALLY INCLUDED IN THE
 CC CODING NUCLEOTIDE SEQUENCE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U14572; NOT_ANNOTATED_CDS.
 CC KW Hypothetical protein.
 CC FT DOMAIN 1 97 FRAME-1.
 CC FT DOMAIN 101 196 FRAME-2.
 CC FT DOMAIN 200 295 FRAME-3.
 CC FT DOMAIN 299 395 FRAME-4.
 CC FT DOMAIN 399 494 FRAME-5.
 CC FT DOMAIN 498 593 FRAME-6.
 CC SQ SEQUENCE 593 AA; 64603 MW; 136EF344AAC012A2 CRC64;
 CC
 CC Query Match 17.6%; Score 357.5; DB 1; Length 593;
 CC Best Local Similarity 48.2%; Pred. No. 2e-24;
 CC Matches 95; Conservative 6; Mismatches 47; Indels 49; Gaps 8;
 CC
 CC OY 4 SLLPRLCNGCAISHNRRLRPGSSDPSASPVAGTGMCHARLIYFLVEMEFHY 63
 CC DB 405 SLLPRLCNGCAISHNRRLRPGSSDPSASPVAGTGMCHARLI-FVLEVTGFLHY 463
 CC OY 64 GQAGLEPTSDPSVSASQASARYRTG--HHARLCIANFCGRNRYSLMCPMS- 113
 CC DB 464 GQAGLEPTSDPSVSASQASAGI-TGVSHRAR---XXXFFEMEPSCCPGMSAMARSLT 519
 CC OY 114 -----PELKOSTCLSLPKCMYRRAAVGGLTLPFLRHRCPTLIDDEVQW 158
 CC DB 520 ATASARVQAILLPQPEX-----LGLQAC-----ATTPGFLYEXXRGFSMLVR----- 564
 CC OY 159 CDHSLQSTPEIKHP 175
 CC DB 565 -----LVNSRQVIRPP 577
 CC
 CC RESULT 7
 CC ALU4_HUMAN
 CC ID ALU4_HUMAN STANDARD. PRT: 603 AA.
 CC AC P39191.
 CC DT 01-FEB-1995 (Rel. 31, Created)
 CC DT 01-FEB-1995 (Rel. 31, Last annotation update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE ALU subfamily SB2 sequence contamination warning entry.
 CC OS Homo sapiens (Human)
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=95021758; PubMed=7935834;
 CC RA Claverie J.-M., Makalowski W.;
 CC "Alu alert.";

RL Nature 371:752-752(1994).
 RN [2]
 RN CONCEPT
 RP MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 RT potentially misleading protein sequences.";
 RL Genomics 12:838-841(1992).
 RN [3]
 RN ALU FAMILIES CLASSIFICATION.
 RP MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation
 RT closely connected with primate lineage history.";
 RL J. Mol. Evol. 27:194-202(1988).
 RN [4]
 RN ALU FAMILIES CLASSIFICATION.
 RP MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes.";
 RL J. Mol. Evol. 32:105-121(1991).
 CC -1- MISCELLANEOUS: VARIOUS ANALYSES (SEE REF.3 AND REF.4) INDICATE
 CC THAT ALU REPEATS FALL INTO 8 SUBFAMILIES. THEREFORE, 8 ALU WARNING
 CC CONSENSUS SEQUENCES HAVE BEEN CONSTITUTED THAT CONTAIN ALL SIX
 CC FRAMES CONCEPTUAL TRANSLATIONS OF EACH OF THESE CLASSES OF ALU
 CC REPEATS.
 CC -1- MISCELLANEOUS: ISOLATED 'X' INDICATES THE PRESENCE OF A STOP
 CC CODON. 'XXX' IS USED TO SEPARATE THE VARIOUS TRANSLATION PHASES.
 CC -1- CAUTION: THIS ALU ENTRY IS PROVIDED IN ORDER TO AVOID THE FURTHER
 CC POLLUTION OF PROTEIN SEQUENCE DATABASES WITH ALU-DERIVED AMINO
 CC ACID SEQUENCES.
 CC -1- CAUTION: ALU REPETITIVE SEQUENCES ARE INTERSPERSED IN HUMAN AND
 CC PRIMATE GENOMES WITH AN AVERAGE SPACING OF 4 KB. SOME OF THEM ARE
 CC ACTIVELY TRANSCRIBED BY POL. III. NORMAL TRANSCRIPTS MAY CONTAIN
 CC CDNA LIBRARIES ALSO CONTAIN PARTIAL AND/OR REARRANGED CNAS
 CC LIGATED WITH ALU-DERIVED SEQUENCE IN ANY ORIENTATION. ALTHOUGH ALU
 CC ELEMENTS (ESPECIALLY SITUATED ON THE COMPLEMENTARY STRAND) HAVE A
 CC GREAT POTENTIAL TO CREATE ADDITIONAL/ALTERNATIVE EXONS,
 CC CONSIDERATION SHOULD BE GIVEN TO THE POSSIBILITY THAT THE PRESENCE
 CC OF AN ALU IN AN OPEN READING FRAME MAY HAVE RESULTED FROM A
 CC CLONING ARTIFACT OR MAY BE DUE TO MISINTERPRETATION OF SEQUENCING
 CC DATA. THIS POINT HAS BEEN OVERLOOKED ON SEVERAL OCCASIONS, WITH
 CC THE CONSEQUENCE OF ERRONEOUS ALU-DERIVED AMINO ACID SEQUENCES
 CC BEING REPORTED.
 CC -1- CAUTION: ANY SIGNIFICANT SIMILARITY OF A PUTATIVE PROTEIN SEQUENCE
 CC WITH AN ALU-TRANSLATED ENTRY MUST BE TAKEN AS A WARNING THAT A
 CC PART OF ALU REPEAT MAY HAVE BEEN ARTIFACTUALLY INCLUDED IN THE
 CC CODING NUCLEOTIDE SEQUENCE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U14570; NOT_ANNOTATED_CDS.
 CC KW Hypothetical protein.
 CC FT DOMAIN 1 98 FRAME-1.
 CC FT DOMAIN 102 199 FRAME-2.
 CC FT DOMAIN 202 300 FRAME-3.
 CC FT DOMAIN 304 401 FRAME-4.
 CC FT DOMAIN 405 502 FRAME-5.
 CC FT DOMAIN 506 603 FRAME-6.
 CC SQ SEQUENCE 603 AA; 65272 MW; B8AADAD46BEA114 CRC64;
 CC
 CC Query Match 17.5%; Score 356.5; DB 1; Length 603;
 CC Best Local Similarity 30.3%; Pred. No. 2.5e-24;
 CC Matches 112; Conservative 28; Mismatches 115; Indels 115; Gaps 14;
 CC
 CC OY 76 PSVASQASARYRTGHHARLCIANFCGRNRYSLMCPMSPELKOSTCLSLPKCMYRRAAV 135

```

Db 156 PSYMGGRMAXTREAELAVSRDCA---TAVRSPAMATERDSVS-----KXXX 202
Qy 136 PGLFLFLRHRG-PTLTODEVQMC-----DHSLQSTP-----169
Db 203 PQA-----VAHACNSTLGGRCGWMRGDRDHPXQGETPSLTIKISRMRAPVVP 257
Qy 170 -----EIKHP-----PASASOYAGTKDMHYTLIFIFLNF 203
Db 258 ATGEAEAEEMREPKRSIQXAEIAPLOSANVRGSRSETPSKKXXF-----FLR 307
Qy 204 QS--LNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSMDYR--PPRLANFVFLVEMG 260
Db 308 RSLASPPRDCGLQWRNLGSLQAPLPDFTFSCSLSPSSMDYRPPRANFLYFXRRG 367
Qy 261 FTMFARLLISGPCDLPASASOSAGITGVSHARLINF-----CLFEMS 306
Db 368 FTLLARMVSIKXPHDPASASOSAGITGVSHARXXFXDGVSLCRPGRTADCSGALSA 427
Qy 307 HSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSMDYR--PPRLANFVFLVEMG 364
Db 428 HCKLRF-----PCSRHSPASASPVACTTG--AHRAHLLFCIFSRDGVSPC 471
Qy 365 LSGMSQTPDL 374
Db 472 XPGMSRSPDL 481

RESULT 8
ALU7_HUMAN
ID ALU7_HUMAN STANDARD: PRT: 593 AA.
AC P39194;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
OS Alu subfamily SQ sequence contamination warning entry.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95021758; PubMed=7935834;
RA Claverie J.-M.; Makalowski W.;
RT "Alu alert.";
RL Nature 371:752-752(1994).
RN [2]
RP CONCEPT.
RX MEDLINE=92241891; PubMed=1572661;
RA Claverie J.-M.;
RT "Identifying coding exons by similarity search: alu-derived and other
RT potentially misleading protein sequences.";
RL Genomics 12:838-841(1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=88333009; PubMed=3138422;
RA Quentin Y.;
RT "The Alu family developed through successive waves of fixation
RT closely connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202(1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=91178815; PubMed=1706781;
RA Jurka J.; Milosavljevic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121(1991).
CC -1- MISCELLANEOUS: VARIOUS ANALYSES (SEE REF. 3 AND REF. 4) INDICATE
CC THAT ALU REPEATS FALL INTO 8 SUBFAMILIES. THEREFORE, 8 ALU WARNING
CC CONSENSUS SEQUENCES HAVE BEEN CONSTITUTED THAT CONTAIN ALL SIX
CC FRAMES CONCEPTUAL TRANSLATIONS OF EACH OF THESE CLASSES OF ALU
CC REPEATS.
CC -1- MISCELLANEOUS: ISOLATED 'X' INDICATES THE PRESENCE OF A STOP
CC CODON, 'xxx' IS USED TO SEPARATE THE VARIOUS TRANSLATION PHASES.

```

```

CC -1- CAUTION: THIS ALU ENTRY IS PROVIDED IN ORDER TO AVOID THE FURTHER
CC POLLUTION OF PROTEIN SEQUENCE DATABASES WITH ALU-DERIVED AMINO
CC ACID SEQUENCES.
CC -1- CAUTION: ALU REPETITIVE SEQUENCES ARE INTERSPERSED IN HUMAN AND
CC PRIMATE GENOMES WITH AN AVERAGE SPACING OF 4 KB. SOME OF THEM ARE
CC ACTIVELY TRANSCRIBED BY POL III. NORMAL TRANSCRIPTS MAY CONTAIN
CC ALU-DERIVED SEQUENCES IN 5' OR 3' UNTRANSLATED REGIONS. HOWEVER,
CC CDNA LIBRARIES ALSO CONTAIN PARTIAL AND/OR REARRANGED CDNAS
CC LIGATED WITH ALU-DERIVED SEQUENCE IN AN ORIENTATION ALTHOUGH ALU
CC ELEMENTS (ESPECIALLY SITUATED ON THE COMPLEMENTARY STRAND) HAVE A
CC GREAT POTENTIAL TO CREATE ADDITIONAL/ALTERNATIVE EXONS,
CC CONSIDERATION SHOULD BE GIVEN TO THE POSSIBILITY THAT THE PRESENCE
CC OF AN ALU IN AN OPEN READING FRAME MAY HAVE RESULTED FROM A
CC CLONING ARTIFACT OR MAY BE DUE TO MISINTERPRETATION OF SEQUENCING
CC DATA. THIS POINT HAS BEEN OVERLOOKED ON SEVERAL OCCASIONS, WITH
CC THE CONSEQUENCE OF ERRONEOUS ALU-DERIVED AMINO ACID SEQUENCES
CC BEING REPORTED.
CC -1- CAUTION: ANY SIGNIFICANT SIMILARITY OF A PUTATIVE PROTEIN SEQUENCE
CC WITH AN ALU REPEAT ENTRY MUST BE TAKEN AS A WARNING THAT A
CC PART OF ALU REPEAT MAY HAVE BEEN ARTIFACTUALLY INCLUDED IN THE
CC CODING NUCLEOTIDE SEQUENCE.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U14573; -; NOT_ANNOTATED_CDSS.
CC DR Hypothetical protein.
CC KW DOMAIN 1
CC FT DOMAIN 101 196 FRAME-1.
CC FT DOMAIN 200 295 FRAME-2.
CC FT DOMAIN 299 395 FRAME-3.
CC FT DOMAIN 399 494 FRAME-4.
CC FT DOMAIN 498 593 FRAME-5.
CC FT DOMAIN 593 644 FRAME-6.
CC SQ SEQUENCE 593 AA; 64417 MW; 54A4F50F3JA6089F CRC64;

Query Match 16.2%; Score 330.5; DB 1; Length 593;
Best Local Similarity 45.9%; Pred. No. 5.2e-22;
Matches 89; Conservative 9; Mismatches 53; Indels 43; Gaps 7;

Qy 4 SLLPRLFCNGAISAHRNRLRPGSSDSPASAVAGITGCMCHARLIYFLVEMEPFLHV 63
Db 405 SLLPRLFCNGAISAHRNRLRPGSSDSPASAVAGITGCMCHARLIYFLVETGFFHHV 463
Qy 64 GQAGLEPTSDPSVSASOSARVYRTG--HHARLCLANFCGRNRVSLMCPMS----- 113
Db 464 GQAGLELLTSGDPPASASOSAGI-TGVSHRAR---XXXFEFERFSCCGMSAVARSRLT 519
Qy 114 -----PELKOSTCTCLSPKCMYRRRAVAGLFLFLRHRCPPLTODEVQWCDH 161
Db 520 ATASRVOAILLPDPEXGLQAP-----ATTPGXFLYFXRRGFTMLAR----- 564
Qy 162 SLSQSTPEIKHP 175
Db 565 -LVNSNXPOVTHP 577

RESULT 9
NEK4_HUMAN
ID NEK4_HUMAN STANDARD: PRT: 841 AA.
AC P51957;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase NEK4 (EC 2.7.1.37) (NIMA-related
DE protein kinase 4) (Serine/threonine protein kinase 2)
DE (Serine/threonine-protein kinase NRK2).
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=9426838; PubMed=8208544;
RA Levedakou E.N., He M., Baptist E.W., Craven R.J., Cance W.G.,
RA Weissh P.L., Simmons A., Naylor S.L., Leach R.J., Lewis T.B.,
RA Bowcock A., Liu E.T.;
RT "Two novel human serine/threonine kinases with homologies to the cell
RT cycle regulating Xenopus Mo15, and NIMA kinases: cloning and
RT characterization of their expression pattern.";
RL Oncogene 9:1977-1988(1994).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBCELLULAR LOCATION: Nucleus (Probable).
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADULT HEART, FOLLOWED BY
CC PANCREAS, SKELETAL MUSCLE, BRAIN, LIVER, KIDNEY, LUNG AND
CC PLACENTA. PRESENT IN MOST PRIMARY CARCINOMAS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIMA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L20321; AAA36568.1;
DR HSSP: 063450; 1A06.
DR Genew: HGNC:11399; NEK4.
DR MIM: 601959;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation.
FT DOMAIN 6 261 PROTEIN KINASE.
FT NP_BIND 12 20 ATP (BY SIMILARITY).
FT BINDING 35 35 ATP (BY SIMILARITY).
FT ACT_SITE 131 131 BY SIMILARITY.
FT MOD_RES 165 165 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 841 AA; 94571 MW; 0DD31920DEFEA58 CRC64;

Query Match 7.9%; Score 161; DB 1; Length 841;
Best Local Similarity 46.2%; Pred. No. 1.1e-06;
Matches 42; Conservative 14; Mismatches 25; Indels 10; Gaps 4;

OY 6 LLPRLENGAISAHNRNLRLPGSSDSPASAPVAGITGCTHAR-LILYFLVEEFLHYG 64
DB 460 LSPRLTECGTILASHNLRLTGSSDSPASAPVAGITGCTHAR-LILYFLVEEFLHYG 64
OY 65 QAGLELP---TSDDPSVAS--QSARRTGH 90
DB 516 RHPDLOPHNGSSEPSLSRORRORRROTET 90

RESULT 10
SCGE HUMAN
ID SCGE HUMAN STANDARD; PRT; 438 AA.
AC 043556;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epsilon-sarcoglycan precursor (Epsilon-Sg).
GN SCGE OR ESG.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Nigro V.;
RT Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98133310; PubMed=9475163;
RA McNally E.M., Ly C.T., Kunkel L.M.;
RT "Human epsilon-sarcoglycan is highly related to alpha-sarcoglycan
RT (adhalin), the limb girdle muscular dystrophy 2D gene.";
RL FEBS Lett. 422:27-32(1998).
RL [3]
RP SEQUENCE OF 269-413 FROM N.A.
RX MEDLINE=98070432; PubMed=9405466;
RA Etlinger A.J., Feng G., Sane J.R.;
RT "Epsilon-sarcoglycan, a broadly expressed homologue of the gene
RT mutated in limb-girdle muscular dystrophy 2D.";
RL J. Biol. Chem. 272:32534-32538(1997).
RL [4]
RP ERRATUM.
RA Etlinger A.J., Feng G., Sane J.R.;
RL J. Biol. Chem. 273:19922-19922(1998).
CC -1- FUNCTION: COMPONENT OF THE SARCOGLYCAN COMPLEX, A SUBCOMPLEX OF
CC THE DYSTROPHIN-GLYCOPROTEIN COMPLEX WHICH FORMS A LINK BETWEEN THE
CC F-ACTIN CYTOSKELETON AND THE EXTRACELLULAR MATRIX.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Sarcolemma
CC (potential).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SCGE-1 (SHOWN HERE) AND SCGE-2;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: UBQUITOUS.
CC -1- SIMILARITY: BELONGS TO THE SARCOGLYCAN ALPHA/EPSILON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF031920; AAC14021.1; ALT_INIT.
DR EMBL: AF036364; AAC04368.1;
DR EMBL: AJ000534; CA004167.1;
DR Genew: HGNC:10808; SCGE.
DR MIM: 604149;
DR CYTOSKELETON; Transmembrane; Glycoprotein; Signal;
KW Alternative splicing
FT CHAIN 23 438 POTENTIAL.
FT DOMAIN 23 293 EPSILON-SARCOGLYCAN.
FT TRANSMEM 294 314 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 315 438 POTENTIAL.
FT DOMAIN 211 317 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 176 176 CYS-RICH.
FT VARSPIC 395 419 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 438 AA; 49754 MW; D87D1899E9C16F23 CRC64;

Query Match 5.0%; Score 101.5; DB 1; Length 438;
Best Local Similarity 29.0%; Pred. No. 0.11;
Matches 29; Conservative 7; Mismatches 25; Indels 39; Gaps 3;

OY 155 EYVQCDHSLQSPTEIKH-----PPASASQVAGTKD--M 187
DB 331 DIQLVHNSAIQKSTLRDMSKNREIAMPSTLPVPHVYTGRIIPRLHTDWDSTNMLM 390
OY 188 HHYTWLFIFFINFLRQSLNSTYQAGVQWRNIGSLQPLRP 227

DB 391 QTCQM-----SFAPVAGVQWRDLGSLQPPPP 418

RESULT 11

MT01_HUMAN STANDARD: PRT: 717 AA.

AC Q9Y222; Q8W257; Q9BS06; Q96FE6;

DT 30-MAY-2000 (Rel. 39, Created)

DR 15-JUN-2002 (Rel. 41, Last sequence update)

DE MT01 protein homolog (Protein CGI-02).

GN MT01.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 5).

RX MEDLINE=20272150; PubMed=10810093;

RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;

RT "Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics."

RL Genome Res. 10:703-713(2000).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 5).

RA Li R.H., Li X.M., Zhang X., Guan M.X.;

RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A. (ISOFORMS 5 AND 6).

RC TISSUE=Lymph, and Placenta;

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- ALTERNATIVE PRODUCTS: 5 isoforms: 1, 2, 3/4 (shown here), 5 and 6; are produced by alternative splicing.

CC -1- SIMILARITY: BELONGS TO THE GIDA FAMILY.

CC -1- CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 9.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF132937; AAD2712.1; ALT_FRAME.

DR EMBL: AF469110; AAL82394.1; -.

DR EMBL: AF469111; AAL82395.1; -.

DR EMBL: AF442963; AAL35894.1; -.

DR EMBL: AF319422; AAG42814.3; -.

DR EMBL: AY078986; AAL85491.1; -.

DR EMBL: AY078985; AAL85490.1; -.

DR EMBL: BC005808; AAL805808.1; -.

DR EMBL: BC011051; AAL11051.1; ALT_INIT.

DR InterPro: IPR002218; GIDA.

DR InterPro: IPR004416; GIDA_sub.

DR Pfam: PF01134; GIDA.1.

DR ProDom: PD003738; GIDA.1.

DR TIGRfam: TIGR00136; gida.1.

DR PROSITE: PS01280; GIDA_1; 1.

DR PROSITE: PS01281; GIDA_2; 1.

DR Alternative splicing.

DR VARSPLIC 1 74 MISSING (IN ISOFORM 6).

FT VARSPLIC 179 275 MISSING (IN ISOFORM 1).

FT VARSPLIC 314 717 MISSING (IN ISOFORM 2).

FT VARSPLIC 376 400 MISSING (IN ISOFORM 1 AND ISOFORM 5).

FT CONFLICT 300 300 H -> Q (IN REF. 1 AND 2; AAL82394/AAL82395).

FT CONFLICT 586 587 LA -> CT (IN REF. 2; AAL35894).

FT SEQUENCE 717 AA: 79963 MW: 78F84DB833BCED3 CR664;

Query Match 5.0%; Score 101; DB 1; Length 717;

Best Local Similarity 86.4%; Pred. No. 0.22;

Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 LLLPRLECGAISAHRLRLPG 26

|||||

DB 381 LLLPRMECGAISAHHLPLPG 402

RESULT 12

2195_HUMAN STANDARD: PRT: 629 AA.

ID 2195_HUMAN

AC 014628;

DT 15-JUL-1998 (Rel. 36, Created)

DR 15-JUL-1998 (Rel. 36, Last sequence update)

DE 15-OCT-2001 (Rel. 40, Last annotation update)

GN Zinc finger protein 195.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.

RX MEDLINE=98008942; PubMed=9344677;

RA Hussey D.J., Parker N.J., Hussey N.D., Little P.F.R., Dobrovic A.;

RT "Characterization of a KRAB family zinc finger gene, ZNF195, mapping to chromosome band 11p15.5."

RL Genomics 45:451-455(1997).

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A (SHOWN HERE), B AND C; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT HEART, BRAIN, PLACENTA, SKELETAL MUSCLE AND PANCREAS, AND IN FETAL LUNG, KIDNEY AND BRAIN.

CC THERE IS LITTLE EXPRESSION IN ADULT LUNG, KIDNEY AND BRAIN.

CC -1- SIMILARITY: BELONGS TO THE KRUPEPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.

CC -1- SIMILARITY: CONTAINS 1 KRAB DOMAIN.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF003540; AAB86596.1; -.

DR HSSP: P08046; IALG.

DR GeneW: HGNC:12986; ZNF195.

DR MIM: 602187; -.

DR InterPro: IPR001909; KRAB.

DR InterPro: IPR000822; znf_C2H2.

DR Pfam: PF00096; zf_C2H2; 9.

DR Pfam: PF01352; KRAB; 1.

DR PRINTS: PRO0048; ZINC_FINGER.

DR ProDom: PD000003; znf_C2H2; 9.

DR SMART: SM00349; KRAB; 1.

DR SMART: SM00355; znf_C2H2; 9.

DR PROSITE: PS50805; KRAB; 1.

DR PROSITE: PS50028; ZINC_FINGER_C2H2_1; 9.

DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 10.

DR Zinc-finger: Metal-binding; DNA-binding; Nuclear protein;

DR Alternative splicing.

DR DOMAIN 4 75 KRAB.

FT DOMAIN 76 243 SPACER.

FT ZN_FING 244 266 C2H2-TYPE.

FT ZN_FING 410 432 C2H2-TYPE.

FT ZN_FING 438 460 C2H2-TYPE.

FT ZN_FING 466 488 C2H2-TYPE.

FT ZN_FING 494 516 C2H2-TYPE.

FT ZN_FING 522 544 C2H2-TYPE.

FT ZN_FING 550 572 C2H2-TYPE.

FT ZN_FING 578 600 C2H2-TYPE.

```

FT 2N.FING 606 628 C2H2-TYPE
SQ SEQUENCE 629 AA: 72303 MM: D9C333FE6DBE7FE CRC64:
Query Match 4.9%; Score 100; DB 1; Length 629;
Best Local Similarity 41.0%; Pred. No. 0.23;
Matches 34; Conservative 8; Mismatches 31; Indels 10; Gaps 5;

OY 216 WRNIGS--LQPLPGFKLFC--PSLLSWPYRRPRLANFVLEVGFTMFARLIL-1S 271
DB 37 YRNLFESVGLTVCKCG--LITLEORKEPWNVKROEADGH-----PEMGFHHATQACIELL 90
OY 272 GPCDLPASASOSAGITGVSHHAR 294
DB 91 GSSDLPASASOSAGITGVNHRAR 113

RESULT 13
RED1_HUMAN STANDARD; PRT; 741 AA.
AC P78563: P78555: 000691: 000692: 000395: 000465:
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Double-stranded RNA-specific editase 1 (EC 3.5.-.-) (dsRNA adenosine
deaminase) (RNA editing enzyme 1).
GN ADAR1 OR RED1 OR DRADA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Gerber A., O'Connell M.A., Keller W.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA Miltaz L., Scott H.S., Rossier C., Antonarakis S.E.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=97265373; PubMed=9111310;
RA Lai F., Chen C.X., Carter K.C., Nishikura K.;
RT "Editing of glutamate receptor B subunit ion channel RNAs by four
alternatively spliced DRADA2 double-stranded RNA adenosine
deaminases."
RL Mol. Cell. Biol. 17:2413-2424(1997).
RN [4]
RN SEQUENCE FROM N.A.
RA Villard L., Tassone F., Haymowicz M., Welborn R., Gardiner K.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RP MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
Park H.-S., Toyoda A., Ishii K., Tokoki Y., Choi D.-K., Soeda E.,
Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
Reischwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,
Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
Shitani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
Mitsushima S., Shimizu N., Nordisk G., Horstischer K., Brandt P.,
Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
Ramer J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
Lehach H., Reinhardt R., Yaspo M.-L.;
RA "The DNA sequence of human chromosome 21."
RL Nature 405:311-319(2000).
CC -!- FUNCTION: EDITING OF THE MESSENGER RNAs FOR GLUTAMATE RECEPTOR
(GLUR) SUBUNITS BY SITE-SELECTIVE ADENOSINE DEAMINATION. EDITS
BOTH THE GLUR-B Q/R AND R/G SITES EFFICIENTLY BUT CONVERTS THE
ADENOSINE IN HOTSPOT1 MUCH LESS EFFICIENTLY.

```

```

CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS: RED1-L/DRADA2B
CC (SHOWN HERE), RED1-S/DRADA2A AND DRADA2C; ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE ADENOSINE DEAMINASE EDITASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 DREM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
CC EMBL: U82120; AAB61686.1; -
CC EMBL: U82121; AAB61687.1; -
CC EMBL: X99227; CA67611.1; -
CC EMBL: X99383; CA67762.1; -
CC EMBL: U76420; AAC51240.1; -
CC EMBL: U76421; AAC51241.1; -
CC EMBL: U76422; AAC51242.1; -
CC EMBL: AF001042; AAB58300.1; -
CC EMBL: AF163301; CAB90493.1; -
CC HSSP: Q91836; 1D12.
CC Genew: HGNC:226; ADAR1.
CC MIM: 601218; -
CC Interpro: IPR002466; A.deamin.
CC Interpro: IPR001159; DS_RBD.
CC Pfam: PF00035; dsrm; 2.
CC Pfam: PF02137; A.deamin; 1.
CC SMART: SM00358; DSRM; 2.
CC PROSITE: PS50141; A.DEAMIN_EDITASE; 1.
CC PROSITE: PS50137; DS_RBD; 2.
CC DR mRNA processing; Hydrolase; 2.
CC KW Alternative splicing; Zinc; RNA-binding; Repeat;
KW Domain 1.
KW Domain 2.
FT DOMAIN 230 300 ZINC (BY SIMILARITY).
FT METAL 394 394 BY SIMILARITY.
FT ACT_SITE 396 396 ZINC (BY SIMILARITY).
FT METAL 451 451 ZINC (BY SIMILARITY).
FT METAL 556 556 MISSING (IN ISOFORM DRADA2A).
FT VARSPLIC 466 505 ARLETFATKAGLGAMVKEPQDOFSILTP -> VH (IN
FT VARSPLIC 713 741 ISOFORM DRADA2C).
FT CONFLICT 30 30 G -> A (IN REF. 4).
FT CONFLICT 423 423 R -> E (IN REF. 4).
FT CONFLICT 475 475 V -> L (IN REF. 4).
SQ SEQUENCE 741 AA: 80763 MM: 02B583414D59C20 CRC64:

Query Match 4.8%; Score 98.5; DB 1; Length 741;
Best Local Similarity 65.7%; Pred. No. 0.37;
Matches 23; Conservative 1; Mismatches 6; Indels 5; Gaps 1;

OY 306 SHSYTQAGVQWPNLGSLOPLPGKRFSCSLPSS 340
DB 467 SRSYTGAGVQWQNHHSLOLPREPGL-----LSDPST 496

RESULT 14
GG2_HUMAN STANDARD; PRT; 501 AA.
ID GYG2_HUMAN
AC G15488: O15489: O15490: O15486: O15485: O15487:
DT 30-MAY-2000 (rel. 39, Created)
DT 30-MAY-2000 (rel. 39, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Glycogenin-2 (EC 2.4.1.186) (GN2).
GN GYG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

```


RC TISSUE: liver;
 RX MEDLINE=98010589; PubMed=9346895;
 RA Mu J., Skurat A.V., Roach P.J.;
 RT "glycogenin-2, a novel self-glucosylating protein involved in liver
 RT glycogen biosynthesis.";
 RL J. Biol. Chem. 272:27589-27597(1997).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20184741; PubMed=10721716;
 RA Zhai L., Mu J., Zong H., DePaoli-Roach A.A., Roach P.J.;
 RT "Structure and chromosomal localization of the human glycogenin-2 gene
 RT GYG2.";
 RL Gene 242:229-235(2000).
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORM BETA).
 RC TISSUE: SKIN;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN CHARACTERIZATION, AND CARBOHYDRATE-LINKAGE SITE TYR-228.
 RX MEDLINE=99074257; PubMed=9857012;
 RA Mu J., Roach P.J.;
 RT "Characterization of human glycogenin-2, a self-glucosylating
 RT initiator of liver glycogen metabolism";
 RL J. Biol. Chem. 273:34850-34856(1998).
 RN [5]
 RN MASS SPECTROMETRY.
 RC TISSUE-Breast cancer;
 RX MEDLINE=21829512; PubMed=11840567;
 RA Harris R.A., Yang A., Stein R.C., Lucy K., Brusten L., Herath A.,
 RA Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J.,
 RA Zvelebil M.J.;
 RT "Cluster analysis of an extensive human breast cancer cell line
 RT proteomics 2:212-223(2002).
 RL [1]
 RL FUNCTION: SELF-GLUCOSYLATES, VIA AN INTER-SUBUNIT MECHANISM, TO
 RL FORM AN OLIGOSACCHARIDE PRIMER THAT SERVES AS SUBSTRATE FOR
 RL GLYCOSYL SYNTHASE.
 CC CATALYTIC ACTIVITY: UDP-glucose + glycogenin = UDP +
 CC glucosylglycogenin.
 CC COFACTOR: SELF-GLUCOSYLATION IS DEPENDENT ON THE PRESENCE OF
 CC DIVALENT METAL IONS OF WHICH MANGANESE ION IS THE MOST EFFECTIVE.
 CC PATHWAY: Glycogen biosynthesis.
 CC SUBUNIT: HOMODIMER, TIGHTLY COMPLEXED TO GLYCOGEN SYNTHASE.
 CC ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS; ALPHA (SHOWN HERE),
 CC BETA, GAMMA, DELTA, EPSILON AND ZETA; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC TISSUE SPECIFICITY: EXPRESSED PREFERENTIALLY IN LIVER, HEART, AND
 CC PANCREAS.
 CC PTM: SELF-GLUCOSYLATED BY THE TRANSFER OF GLUCOSE RESIDUES FROM
 CC UDP-GLUCOSE TO ITSELF, FORMING AN ALPHA-1,4-GLUCAN OF AROUND 10
 CC RESIDUES ATTACHED TO TYR-228.
 CC MASS SPECTROMETRY: MW=55211.89; METHOD-MALDI.
 CC SIMILARITY: BELONGS TO THE GLYCOGENIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U94363; AAB84377.1; -
 DR EMBL: U94363; AAB84378.1; -
 DR EMBL: U94364; AAB84379.1; -
 DR EMBL: U94357; AAB84373.1; -
 DR EMBL: U94358; AAB84374.1; -
 DR EMBL: U94360; AAB84375.1; -
 DR EMBL: U94361; AAB84376.1; -
 DR EMBL: A179624; AAB61855.1; -
 DR EMBL: A179615; AAB61855.1; JOINED.
 DR EMBL: A179616; AAB61855.1; JOINED.

DR EMBL: A179617; AAB61855.1; JOINED.
 DR EMBL: A179618; AAB61855.1; JOINED.
 DR EMBL: A179619; AAB61855.1; JOINED.
 DR EMBL: A179620; AAB61855.1; JOINED.
 DR EMBL: A179621; AAB61855.1; JOINED.
 DR EMBL: A179622; AAB61855.1; JOINED.
 DR EMBL: A179623; AAB61855.1; JOINED.
 DR EMBL: BC023152; AAB61855.1; JOINED.
 DR Gene: HGNC:4700; GYG2.
 DR MIM: 300198; -
 DR InterPro: IPR002495; GT_8.
 DR Pfam: PF01501; Glyco_transf_8; 1.
 DR Transferrase; Glycogen biosynthesis; Alternative splicing;
 KM Glycoprotein; Polymorphism.
 FT ACT_SITE 119 119 BY SIMILARITY.
 FT CARBOHYD 228 228 O-LINKED (GLC...).
 FT VARSPLIC 3 33 MISSING (IN ISOFORM BETA).
 FT VARSPLIC 3 42 MISSING (IN ISOFORM GAMMA).
 FT VARSPLIC 378 448 MISSING (IN ISOFORM DELTA).
 FT VARSPLIC 407 501 MISSING (IN ISOFORM EPSILON).
 FT VARSPLIC 413 448 MISSING (IN ISOFORM ZETA).
 FT VARIANT 270 270 V->A.
 FT MUTAGEN 228 228 /FTID=VAR.010401.
 FT MUTAGEN 230 230 Y->F: LOSS OF ACTIVITY.
 FT CONFLICT 313 313 Y->F: NO LOSS OF ACTIVITY.
 FT CONFLICT 413 413 H->R (IN REF. 3).
 FT CONFLICT 462 464 MISSING (IN REF. 1; AAB84378).
 FT CONFLICT 462 464 EKV->AGI (IN REF. 1; AAB84376).
 SQ SEQUENCE 501 AA: 55211 MW: 28205 FAD5A7657 CRC64:
 Query Match 4.8%; Score 98; DB 1; Length 501;
 Best Local Similarity 19.6%; Pred. No. 0.27;
 Matches 91; Conservative 44; Mismatches 100; Indels 230; Gaps 22;
 QY 57 EMEFLHYGAGLELPISDPSVASOSARTRTHHARCLAN--FC----- 100
 DB 3 ETEFHHAQAGLELLRSNSPTASAGMTVDQAFVLTATNDIYCOGALVIGOSLRH 62
 QY 101 --GRNRSLMCPSMS----- 125
 DB 63 RLTRKLVLTTPVOSSLLRVILSKYFDEVIEVNLIDSADYIHLAFLKRPGLG---LTLP 118
 QY 126 K--CW-----DYRAAVP-----GLFLFLRHR 147
 DB 119 KLKQWTLTHSKCVFLDADTLVLSNDELDFKREFSAAPRPGPCDFNSGVFPQPSLHT 178
 QY 148 CPTLTODEVQWCHSSLOPSTPEIKHPPASOVAGTKDMHHTWLIETFNFLROSLN 207
 DB 179 HKLLQHAM--EHGSEFDGADQGLN---SFFRNWSTTDIHKH---LPIYVLSNTWY 228
 QY 208 SYVQAGVQWNN-----LGSLOPLPRGFLFCPSLSSWDYRRPRRLANFVFLVEMG 260
 DB 229 TYSPAKRFQESSAKVYHFLGSMKP-----WNYKYNPOQS---VLEQG 268
 QY 261 -----FTMFARLIL-----ISGCDLPASASQ 282
 DB 269 SVSSQHQAAFLHLMWTVYQNNVLPYKSVQAGEARASPGHTLCHSVGPRC-----ADS 323
 QY 283 SAGITGVSHHARILFNCLFEMESHVYQAGVQWPN--LGSLOP----- 324
 DB 324 ASGVG-----EPCENSTPSAGVPCANSIPNOAGLPAPTOIVDETLS 368
 QY 325 LPPGLKR-----FSC--LSLPSSMDYGLRPHHPANF 353
 DB 369 LPEGRSEDMIACPETETPAVITCDPLSOPS-----PQADF 405
 RESULT 15
 ID RRPQ_NMV STANDARD; PRT; 1643 AA.
 AC R15095;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)

BT 15-JUN-2002 (Rel. 41, last annotation update)
DE RNA replication protein (186 kDa protein) (ORF 1) [Contains: RNA-
directed RNA polymerase (EC 2.7.7.48); Probable helicase].
OS Narcissus mosaic virus (NMV)
CC Viruses: ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxID=12180;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89279206; PubMed=2732689;
RA Zuidema D., Linthorst H.J.M., Huisman M.J., Asjes C.J., Bol J.F.;
RT "Nucleotide sequence of narcissus mosaic virus RNA."
RU J. Gen. Virol. 70:267-276(1989).
CC -I- FUNCTION: RNA-replication. The central part of this protein
possibly functions as an ATP-binding helicase.
CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA](N).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D13747; BAA02891.1; -
DR PIR: J0470; RRGWNV.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
KW ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;
KW transferase.
FT NP_BIND 868 875 ATP (POTENTIAL).
SQ SEQUENCE 1643 AA; 186304 MW; ADA6EFC0FE2EC43 CRC64;
Query Match 4.7% Score 96.5; DB 1; Length 1643;
Best Local Similarity 20.6% Pred. No. 1.4; Mismatches 113; Indels 95; Gaps 14;
Matches 64; Conservative 39;
QY 100 CGRNRYSLMCPSPSPDLKSTCLSLPKCWYRRAAVPGLFLFLRHRCPTLTODEVQMC 159
DB 869 CGSGCKSFALQEMMSLKEDQSVV-----TVVTPVLLRNDWQTKLPILPADVFKTF 920
QY 160 DHSSLOPSTPEI-----KHPPASQVAGTKDMHHYTLIFLIFNLRQSLNSVTQAG 213
DB 921 EKSVIOPCNPIVDDYTKLPGLIESV---MHQN-VFILLGDNKQSVYHETNPE 975
QY 214 VQMRNIGSLQPLPPGFKLFS--CPSLSSWDYRRPRLANFVFLVEMGFTMFARLLIS 271
DB 976 AV-----IALPENVEIFSPCEFLNA-THRNVKDLN-----KLGV----- 1012
QY 272 GPCDLPASQASAG--ITGVSHHAR-----LIFNCLFEMESHSVTQAGVQMPN 318
DB 1013 -----YSERGKLVNFAHSLKASRIPLMLVPSTMRNAMFDMGHSMYAGQGLT 1064
QY 319 LGSLOPLPPGLKRF-----SCLSL-----PSSWDYGHLPHPANCFIRGC 360
DB 1065 APKIQILLDNHTQFCSERVLYTCLSAVRARIHFINTGPTTGDY-----WAKLE 1112
QY 361 VSPYLSGWSQT 371
DB 1113 STPYLAKAFIDT 1123

Search completed: July 17, 2003, 18:17:53
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 17, 2003, 18:15:09 ; Search time 40 Seconds

(without alignments)
901.260 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MEFSLLPRLCNGAISAH.....FIRGVSPLYSGMSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	382.5	18.8	673	4 F40201	artifact-warning s
2	315.5	15.5	613	4 C40201	artifact-warning s
3	315.5	15.5	627	4 A40201	artifact-warning s
4	291.5	14.3	597	4 E40201	artifact-warning s
5	197.5	9.7	579	4 D40201	artifact-warning s
6	175	8.6	499	2 S65657	alpha-1C-adrenergic
7	167	8.2	46	2 I54375	gene NF2 protein
8	161	7.9	841	1 I78885	serine/threonine-s
9	157	7.7	39	2 I54374	gene NF2 protein
10	151.5	7.4	301	4 B40201	artifact-warning s
11	139.5	6.9	100	2 A46010	X-linked retinopat
12	137.5	6.8	79	2 A56194	thromboxane A-2 re
13	137.5	6.8	407	2 T02670	probable thromboxa
14	128	6.3	53	2 A42442	integrin beta-1 ch
15	112	5.3	440	2 A26359	decay-accelerating
16	106.5	5.2	1125	2 T19193	hypothetical prote
17	105	5.2	522	2 T08711	gamma-adaptin homo
18	102.5	5.0	331	2 S59501	interferon recepto
19	99.5	4.9	331	2 A54295	interferon alpha/b
20	96.5	4.7	1643	1 R86GNV	genome polypeptide
21	96	4.7	458	2 E82175	conserved hypotet
22	92	4.5	964	2 T15342	transcription fact
23	91.5	4.5	542	2 S36008	transcription fact
24	91.5	4.5	631	1 A36749	genome polypeptide
25	91.5	4.5	1776	1 R86PYM	middle surface ant
26	90.5	4.4	282	1 S20752	DNA-directed DNA p
27	90.5	4.4	832	1 SAVLC2	large surface anti
28	90	4.4	431	1 SAVLC2	probable xylogluca
29	89	4.4	299	2 F84785	

30	89	4.4	333	2 A12131	hypothetical prote
31	88	4.4	1162	2 B97852	hypothetical prote
32	88	4.3	209	2 J04244	heat-shock 27k pro
33	88	4.3	726	2 S18208	radphillin-3A-inter
34	88	4.3	832	2 S71785	DNA-directed DNA p
35	87.5	4.3	494	2 S39607	transcription fact
36	87	4.3	226	1 J01574	major surface anti
37	87	4.3	226	2 J02101	surface antigen -
38	87	4.3	470	2 T05258	glycine hydroxymet
39	87	4.3	1153	2 A49676	nitric-oxide synth
40	86.5	4.3	360	2 T45956	hypothetical prote
41	86.5	4.3	1886	2 S04921	nuclear pore prote
42	86	4.2	262	2 AF2291	hypothetical prote
43	86	4.2	625	2 E96721	hypothetical prote
44	85.5	4.2	832	1 J01574	DNA-directed DNA p
45	84.5	4.2	191	2 C72455	hypothetical prote

ALIGNMENTS

RESULT 1									
F40201									
artifact-warning sequence (translated ALU class F) - human									
C:Species: Homo sapiens (man)									
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #ext_change 19-May-2000									
C:Accession: F40201									
R:Claverie, J.M.									
Personal communication, 1992									
A:Reference number: A40201									
A:Accession: F40201									
A:Molecule type: DNA									
A:Residues: 1-673 <CLA>									
R:Claverie, J.M.									
Genomics 12, 838-841, 1992									
A:Title: Identifying coding exons by similarity search: Alu-derived and other potenti									
A:Reference number: A40200; MUID:92241891; PMID:1572661									
A:Contents: annotation									
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames o									
In-frame stop codons are shown as 'X'.									
C:Comment: Any significant similarity of a predicted protein sequence to a portion of									
Query Match									
Best Local Similarity 33.2%; Pred. No. 7.3e-26; Length 673;									
Matches 128; Conservative 19; Mismatches 111; Indels 127; Gaps 17;									
QY	16	ISAHRLRLPGSSDSPASAPVAGTGMCTHARLILYFLVEMEFHVGAGLEPTSD	75						
DB	371	VLAHCSLNLGSSDSPASVSRVAGITGMGRHMLI-YVFLIEQFHVQNGKLTSSD	429						
QY	76	PSVASQSARVYRTGHHARLCLANFCGRNVSLMCPSPSE-LKOSTCLSLPKCMDYRAA	134						
DB	430	-----LPSMSKPYLVGXAXATTPSXXXF-V	453						
QY	135	VGGFLIFLFL-----RHRCF---TLTODEVQWCHSLQPTPEIKHPASA	178						
DB	454	FGGFEFFFPALFLRLXALALPRLPCSGKFWLTAASTSW-QAILPLSPVXGLQAWA	512						
QY	179	SOVAGT---KDMHYTWLIFIFNFIQLQSLSNVTOAGVQRNLGSLQPLPPGFLLFSC	234						
DB	513	A-IPGKFMVFXRHSHTMLI-----RLVNSKXQV-----ICPPG-----	546						
QY	235	PGLSSWDYRR-PRLANFVFLVEMGFTMARLILISGCDLPASASQAGITGVSHHA	293						
DB	547	--LPKCMDDRRPRPHAXXXFLF--GVFF-----	572						
QY	294	RLIFNCLFEMESHVSTQGVQW---PNLGSLQPLPPGKRSCLSPSSNDYGHLPHP	350						
DB	573	--FPLLCFCFXDRPLXHPG--WSAVASFGSLQPPGPKRPSCLCLPSNDYRHGPPFL	628						
QY	351	ANFCIFIRGVSPLYSGMSQTPDLR	375						

Db 629 ANLCIFNRDVTSPCCXGWSOTPDLK 653

RESULT 2

C40201

artifact-warning sequence (translated ALU class C) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: C40201

R:Claverie, J.M.
personal communication, 1992

A:Reference number: A40201

A:Accession: C40201

A:Molecule type: DNA

A:Residues: 1-613 <CL>

R:Claverie, J.M.
Genomics 12, 838-841, 1992

A:Title: Identifying coding exons by similarity search: Alu-derived and other potential

A:Reference number: A40200; MUID:92241891; PMID:1572661

A:Contents: annotation

C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c

in-frame stop codons are shown as 'x'.

C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 15.5%; Score 315.5; DB 4; Length 613;

Best Local Similarity 33.4%; Pred. No. 5.6e-20;

Matches 98; Conservative 9; Mismatches 73; Indels 113; Gaps 8;

QY 6 LLPLRECGAISAHRNLRPGSSDSPASAPVAGITGCTHARLLTFLVEMEFHVGQ 65

DB 424 LSPRLCEGSAISHCKLRPGCHSPASAFVCRGTGARTMPS-XFEVFLVEMGCHVQ 482

QY 66 AGLEPTSDSDSVASQASARYRTGHHARLCLANFCGRNRVSLMCPSPBELKOSTCLSLP 125

DB 483 AGLELLAS-----XSTHLCCLP 498

QY 126 KCMWDRR---AAVGLFLFELRHRCPTLTQDEVQWCHSLQSTPEIKHPASA-- 178

DB 499 KCMWDRRRLPHRPAAXXFFFFP-----XDRSLC-----HFGSMAAR 536

QY 179 SOVAGTKMHHYTWLIFLFTINFLQSLNSVTQACVQWRNIGSLQPLPGRFLFSCPSLL 238

DB 537 SRLTASSTRYVAIL-----PQPSAVVGLQAP-- 565

QY 239 SSMWDRRRLPHRPAAXXFFFFP-----XDRSLC-----HFGSMAAR 291

DB 566 -----PCPANTLYFXKXWGFAMLARLVNSMWPDPPTASOSAGITGVSH 610

RESULT 3

A40201

artifact-warning sequence (translated ALU class A) - human

C:Species: Homo sapiens (man)

C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000

C:Accession: A40201

R:Claverie, J.M.
personal communication, 1992

A:Reference number: A40201

A:Accession: A40201

A:Molecule type: DNA

A:Residues: 1-627 <CL>

R:Claverie, J.M.
Genomics 12, 838-841, 1992

A:Title: Identifying coding exons by similarity search: Alu-derived and other potential

A:Reference number: A40200; MUID:92241891; PMID:1572661

A:Contents: annotation

C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c

in-frame stop codons are shown as 'x'.

C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 15.5%; Score 315.5; DB 4; Length 627;

Best Local Similarity 33.4%; Pred. No. 5.7e-20;

Matches 98; Conservative 13; Mismatches 67; Indels 115; Gaps 7;

QY 6 LLPLRECGAISAHRNLRPGSSDSPASAPVAGITGCTHARLLTFLVEMEFHVGQ 65

DB 436 LSPRLCEGSAISHCKLRPGCHSPASAFVCRGTGARTMPS-XFEVFLVEMGCHVQ 494

QY 66 AGLEPTSDSDSVASQASARYRTGHHARLCLANFCGRNRVSLMCPSPBELKOSTCLSLP 125

DB 495 DGLDLTL-----SXSR-----LCILP 510

QY 126 KCMWDRRAAVPGLFIFLFRHRCPTLTQDEVQWCHSLQSTPEIKHPASASQVATK 185

DB 511 KCMWDRRETAPG----- 522

QY 186 DMHHTWLIFFIF-----NFLQSLNSVTQ---AGVQWRNIGSLQPLPGRFLFSCPS 236

DB 523 ---XXFFLFFVIFFRDQVSLCROGWSAVARSLPASASVHAILLPQPKXLGLQAP 579

QY 237 LSSMDYRRPRLPHRPAAXXFFFFP-----XDRSLC-----HFGSMAAR 289

DB 580 L-----RPAFLYFXKXRGFTVAVRNVISXPRDPALASOSAGITGV 622

RESULT 4

E40201

artifact-warning sequence (translated ALU class E) - human

C:Species: Homo sapiens (man)

C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000

C:Accession: E40201

R:Claverie, J.M.
personal communication, 1992

A:Reference number: A40201

A:Accession: E40201

A:Molecule type: DNA

A:Residues: 1-597 <CL>

R:Claverie, J.M.
Genomics 12, 838-841, 1992

A:Title: Identifying coding exons by similarity search: Alu-derived and other potent

A:Reference number: A40200; MUID:92241891; PMID:1572661

A:Contents: annotation

C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames o

in-frame stop codons are shown as 'x'.

C:Comment: Any significant similarity of a predicted protein sequence to a portion of

Query Match 14.3%; Score 291.5; DB 4; Length 597;

Best Local Similarity 30.5%; Pred. No. 7.2e-18;

Matches 118; Conservative 24; Mismatches 106; Indels 139; Gaps 20;

QY 5 LLPLRECGAISAHRNLRPGSSDSPASAPVAGITGCTHARLLTFLVEM 58

DB 310 LLPLREYDGTISP--QTLPEFKQSPVSAFVAGITGCHHQAOLLYFXKXWGFSLV 367

QY 59 EFLHVGQAGLEPLTSDSDSVASQASARYRT-----GHHARLC 95

DB 368 KLVNSQGPQVIRPALNSOSAGIT-GMSHTWXXKXFFPETEFCSCCPBRSTMAOSHRLC 426

QY 96 LANFCGRNRVSLMCPSPBELKOSTCLSPKCMWDRRAAVPGLFIFLFRHRCPTLTQDE 155

DB 427 ---FLGSSNSLSLP--PEX-----LGLQAC-----ATPSXRYIF-----SRDG 461

QY 156 V-----QWCHSLQSTPEIKHPASASQVAGTKDMHHTWLIFFIFNFLQSLNSVT 210

DB 462 VSPCWSWSQTPNLRKXSAFPYT--PVALGLQAXATTP-----GXKXFFLRSSFLA 512

QY 211 QAGVQWRNIGSLQPLPGRFLFSCPSLLSSMDYRR--PRLANFVFLVEMGFTME 264

DB 513 QAGVRMHLTA-----NFASVQAIIISCLSPSSMDYRRAPPRAN-FILVEMGFL-- 563

QY 265 ARLLISGCDLPPASASOSAGITGVSHHARLIFNFCLEFEMESHVQAGVQWRNIGSLQ 324

DB 564 -----HYGQAGLKLPTSGD--- 577


```
QY      66 AGLELPTSDDPVSVA 80
        ||::| :|:|
Db      60 AGVQLLPFEPTTCKA 74
```

RESULT 13
T02670

probable thromboxane A₂ receptor isoform beta - human

C.Species: homo sapiens (man)
C.Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 15-Sep-2000
C.Accession: T02670
R.Lamerdin, J.E.; McCreedy, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.; G.J.; Dangaran, L.; Poundstone, P.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; A.P.; Qian, G.; Krommiller, B.; Arellano, A.; Montgomery, M.; Ow, D.; Nolan, M.
Submitted to the EMBL data library, June 1998
A.Authors: Trong, S.; Kobayashi, A.; Olsen, A.S.; Carrano, A.V.
A.Description: Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a serine
A.Reference number: Z14696
A.Accession: T02670
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-407 <LAMB>
A.Cross-references: EMBL:AC005175; NID:g3253116; PID:g3253117
C.Superfamily: prostaglandin E receptor EPI

Query Match	6.88;	Score 137.5;	DB 2;	Length 407;
Best Local Similarity	49.38;	Pred. No. 0.00019;		
Matches 37; Conservative	8;	Mismatches 25;	Indels 5;	Gaps 2

Dy 6 LIPRECNGLISAHRRLRPSSSDSPASASPAGITGCTHARLLYFLVEMFELHVQ 65

Dd 333 LMPSELEYSGTISAHCNLRLPSSSDSPASASRAAGITGV-SHCARPCLMEDPEFDL---- 383

```
QY      66 AGLELPTSDPSVSA 80
      ||::| :|:|
Db     388 AGVQLLPFEPTGKA 402
```

RESULT 14

Integrin beta-1 chain, splice form beta-1-S - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text-change 20-Aug-1999
C:Accession: A42442
R:Languiño, L.R.; Ruoslahti, E.
J. Biol. Chem. 267, 7116-7120, 1992
A:Title: An alternative form of the integrin beta 1 subunit with a variant cytoplasmic c
A:Reference number: A42442; MUID:92202279; PMID:1551917
A:Accession: A42442
A:Molecule type: mRNA
A:Residues: 1-53 <L>
A:Cross-references: GB:M84237; NID:g186221; PIDN:AAA74403.1; PID:g186223
A:Note: sequence extracted from NCBI backbone (NCBIN:92159, NCBIPI:92160)
C:Genetics:
A:Gene: GDB:ITGB1; FNRB
A:Cross-references: GDB:118732; OMIM:135630
A:Map position: 10p11.2-10p11.2
C:Superfamily: Integrin beta chain; laminin-type EGF-like homology
C:Keywords: alternative splicing; duplication; heterodimer; membrane protein

Query Match	6.38;	Score	128;	DB 2;	Length	53;			
Best Local Similarity	63.28;	Pred. No.	0.00012;						
Matches	24;	Conservative	5;	Mismatches	9;	Indels	0;	Gaps	0

```
QY      306 SHSVTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSWDY 343
      | | | | | : | | | | | : | | | | | : | |
Db      6 SLVAQPGVQWCDISSLOPLTSRFQFSCLSLPSWDY 43
```

RESULT 15
A26359
decay-accelerating factor, splice form 1 precursor - human
N;Alternate names: decay-accelerating factor 5; decay-accelerating factor secreted form

C:Species: Homo sapiens (man)
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 02-Jun-2000
C:Accession: A26359; A39702; S16187; S23138; A27258
R:Casas, I.W.; Daultz, M.A.; Rhee, L.; Weddell, G.; Martin Jr., D.W.; Nussenzweig, V.
Nature 325, 545-549, 1987
A:Title: Cloning of decay-accelerating factor suggests novel use of splicing to generate
A:Reference number: A26359; MUID:87115845; PMID:2433596
A:Accession: A26359

A: Molecule type: protein
A: Residues: 35-47 <BIO>
R: Nakano, Y.; Sumida, K.; Kikuta, N.; Miura, N.H.; Tobe, T.; Tomita, M.
Biochim. Biophys. Acta 1116, 235-240, 1992
A: Title: Complete determination of disulfide bonds localized within the short consensus
A: Reference number: S23138; MUID:92305034; PMID:1377029

A:Status: preliminary
A:Molecule type: protein
A:Residues: 35-41;65-68;79-81;93-103;128-134;143-145;155-159;162-168;188-192;203-204
R:Suñita, Y.; Negro, T.; Matsuda, T.; Sakamoto, T.; Tomita, M.
J Biochem 100:143-150 1986

3: Biochem. 100, 143-50, 1958
 4: Title: Improved method for the isolation and preliminary characterization of human
 5: Reference number: A27258; MUID:87008461; PMID:2428813
 6: Accession: A27258
 7: Molecular type: protein
 8: Residues: 35, 'X', 37, 'G', 39-51, 'P', 53-55, 'X', 57-58, 'X', 60-63 <SU>
 9: Note: Gly-35 and Leu-38 were also found
 10: Comment: For an alternative splice form, see PIR:B26359

F;95/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	5.5%	Score 112;	DB 2;	Length 440;
Best Local Similarity	20.6%	Pred. No. 0.038		
Matches 87; Conservative	43;	Mismatches 100;	Indels 192;	Gaps 22

```

QY      22 LRLPSSDSP--ASASPVAGITGMCTHARLLYFLVEMEPFLHVQAGLELFTSDDPVS 79
       ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      70 WKIPGKDSVICKGQSQMSDIEFCNRS-----CEVPRRLN---S 106

```

Dd 107 ASLKQPIQNY-----FPGVGEVEECRCGYNREESLSBKL--TCLDNLK-W----- 151

Qy 133 AAVGGLLEFLNRKPTLTDEVDKCHSSLDQSTPEIKNRPASASQVASTKDMNNHTW 192

 |:::| | |:::| | |

```

Db      152 -----STAVEFCCKKNC-ENPGEIRN---GQIDVHG-----6 179
QY      193 LIFILEFN-----LROSLNS---VYQAGVOMRNLGSLDPLPBGKLFSCPSLSMSD 242
      180 ILRGATISFCSCNMGYKLFSGTSSFCLLISGSSVQMD-----PLPECREITY-CPA----- 227
Db      243 YRRPPRLANFV-----PLVEMGFTTFARLLILI-----SCP----- 273
      228 ---PQIDNGIIGGERDHGYGRQSVYACNKGFTMIGHSIYCTVNNDEGEWGSQPPPCR 284
Db      274 -----CDLP-----ASASQAGITGVSH-----292
QY      285 GKSLTSKVPEYOKPTTVAVPTTEVXSPSSQKTTTTPTPNAQATRSIPVASTTHHFETT 344
Db      293 -----ARLIFNFCILFEMESHSVYQACVQPNLGSLOPLPBGKLRFSCLSLPSSW 341
Db      345 PNKGSGTTSCTTRLLSG-----SRPVYQAGMRWCDCRSSLOSTRPGCFKRSRFHSPLPSSW 397
QY      342 DY 343
      398 Y 399
Db

```

Search completed: July 17, 2003, 18:20:09
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 17, 2003, 18:16:09 ; Search time 27 Seconds
(without alignments)
408.652 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MEFSLLPRLRLECGAISAHNR.....FIRGVSPYLSGMSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfills1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2034	100.0	375	2	US-08-454-557C-121
2	2034	100.0	375	2	US-08-340-426D-121
3	2034	100.0	375	2	US-08-450-673C-121
4	1415.5	69.6	397	5	PCT-US95-17111A-121
5	250	12.3	132	4	US-09-605-785-573
6	218	10.7	1079	3	US-09-058-489-22
7	210.5	10.3	110	4	US-09-227-357-193
8	186	9.1	56	4	US-09-227-357-577
9	168.5	8.3	500	4	US-09-265-630-13
10	158	7.8	122	4	US-09-227-357-285
11	155	7.6	388	4	US-09-265-630-11
12	141	6.9	90	4	US-09-227-357-171
13	129.5	6.4	99	4	US-09-288-143-168
14	128	6.3	48	3	US-08-951-200A-1
15	122	6.0	53	3	US-08-951-200A-8
16	120	5.9	72	4	US-09-227-357-655
17	109	5.4	579	3	US-08-704-711A-1
18	109	5.4	579	4	US-09-521-220-1
19	99.5	4.9	331	4	US-08-585-191A-2
20	99.5	4.9	331	4	US-08-472-402A-2
21	98.5	4.8	714	4	US-09-347-878-16
22	92.5	4.5	47	4	US-09-227-357-656
23	91.5	4.5	631	4	US-08-927-219-127
24	90.5	4.4	105	4	US-08-556-627A-4
25	89	4.4	32	4	US-09-288-143-171
26	87	4.3	630	4	US-08-927-219-2
27	87	4.3	630	4	US-08-927-219-4

28	87	4.3	1146	4	US-09-126-109-12	Sequence 12, App1
29	83	4.1	716	2	US-08-766-982-1	Sequence 1, App1
30	83	4.1	716	4	US-09-296-219-1	Sequence 1, App1
31	82.5	4.1	226	5	PCT-US96-10602-14	Sequence 14, App1
32	82.5	4.1	493	4	US-08-999-774A-12	Sequence 12, App1
33	82.5	4.1	1306	4	US-08-999-774A-13	Sequence 13, App1
34	82	4.0	544	4	US-09-087-134-14	Sequence 14, App1
35	82	4.0	655	4	US-09-228-986-70	Sequence 70, App1
36	82	4.0	1153	1	US-08-314-917-2	Sequence 2, App1
37	82	4.0	1153	1	US-08-265-046-2	Sequence 2, App1
38	82	4.0	1153	2	US-08-465-522-2	Sequence 2, App1
39	82	4.0	1153	5	PCT-US93-11401-2	Sequence 2, App1
40	82	4.0	1153	5	PCT-US95-07849-2	Sequence 2, App1
41	81.5	4.0	281	1	US-08-105-483-214	Sequence 214, App
42	81.5	4.0	281	1	US-08-709-209-214	Sequence 214, App
43	81.5	4.0	281	1	US-08-458-101-214	Sequence 214, App
44	81.5	4.0	351	5	PCT-US96-10602-4	Sequence 4, App1
45	81.5	4.0	389	1	US-08-105-483-216	Sequence 216, App

ALIGNMENTS

```
RESULT 1
US-08-454-557C-121
; Sequence 121, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Mandis, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-454-557C-121

Query Match      100.0%; Score 2034; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.4e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MEFSLLPRLRLECGAISAHNRNLPGSSDSPASAPYAGTGTCTARLLTFVLYEMER 60
        |||
Db      1 MEFSLLPRLRLECGAISAHNRNLPGSSDSPASAPYAGTGTCTARLLTFVLYEMER 60
QY      61 LHVQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPFELKOST 120
```

Db 61 LHVQAGLELPTSDPSVASQSARVYTGHHARLCLANFCGRNRVSLMCPSPMSELKOST 120
QY 121 CLSLPKCWDYRRAAVPGLIFLFLRHRCPTLLTODEVOMCHSSLOPSTPEIKHPASASQ 180
Db 121 CLSLPKCWDYRRAAVPGLIFLFLRHRCPTLLTODEVOMCHSSLOPSTPEIKHPASASQ 180
QY 181 VAGTKDMHHYTWLFFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLSS 240
Db 181 VAGTKDMHHYTWLFFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLSS 240
QY 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIFNFC 300
Db 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVSVOAGVQWPNLGSLOPLPGKLRFSCLSLPSSWDYGHLPHPANCFIFIRGG 360
Db 301 LFEMESHVSVOAGVQWPNLGSLOPLPGKLRFSCLSLPSSWDYGHLPHPANCFIFIRGG 360
QY 361 VSPYLSGMSQTPDLR 375
Db 361 VSPYLSGMSQTPDLR 375

RESULT 2
US-08-340-426D-121
Sequence 121, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-426D-121

Query Match 100.0%; Score 2034; DB 2: Length 375;
Best Local Similarity 100.0%; Pred. No. 2.4e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEFSLPLRLCEGATSAHRNRLRLPGSSDPASASPVAGITGCTHARLLIFFLVEMEF 60
Db 1 MEFSLPLRLCEGATSAHRNRLRLPGSSDPASASPVAGITGCTHARLLIFFLVEMEF 60

QY 61 LHVQAGLELPTSDPSVASQSARVYTGHHARLCLANFCGRNRVSLMCPSPMSELKOST 120
Db 61 LHVQAGLELPTSDPSVASQSARVYTGHHARLCLANFCGRNRVSLMCPSPMSELKOST 120
QY 121 CLSLPKCWDYRRAAVPGLIFLFLRHRCPTLLTODEVOMCHSSLOPSTPEIKHPASASQ 180
Db 121 CLSLPKCWDYRRAAVPGLIFLFLRHRCPTLLTODEVOMCHSSLOPSTPEIKHPASASQ 180
QY 181 VAGTKDMHHYTWLFFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLSS 240
Db 181 VAGTKDMHHYTWLFFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLSS 240
QY 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIFNFC 300
Db 241 WDYRRPRLANFEVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVSVOAGVQWPNLGSLOPLPGKLRFSCLSLPSSWDYGHLPHPANCFIFIRGG 360
Db 301 LFEMESHVSVOAGVQWPNLGSLOPLPGKLRFSCLSLPSSWDYGHLPHPANCFIFIRGG 360
QY 361 VSPYLSGMSQTPDLR 375
Db 361 VSPYLSGMSQTPDLR 375

RESULT 3
US-08-450-673C-121
Sequence 121, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-450-673C-121

Query Match 100.0%; Score 2034; DB 2: Length 375;
Best Local Similarity 100.0%; Pred. No. 2.4e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEFSLPLRLCEGATSAHRNRLRLPGSSDPASASPVAGITGCTHARLLIFFLVEMEF 60
Db 1 MEFSLPLRLCEGATSAHRNRLRLPGSSDPASASPVAGITGCTHARLLIFFLVEMEF 60

Qy	61	LHVQAGLELPTSDPSVASASQATYRGHNARLCLANCGNRNVSILMCPMSPELKQST	120
Dp	61	LHVQAGLELPTSDPSVASASQATYRGHNARLCLANCGNRNVSILMCPMSPELKQST	120
Qy	121	CLSLPKCDYRRAAPGLFLEFLRHRCPTLQODEVOMCDHSLSLOPSTEIKRPPASASQ	180
Dp	121	CLSLPKCDYRRAAPGLFLEFLRHRCPTLQODEVOMCDHSLSLOPSTEIKRPPASASQ	180
Qy	181	VAGTRDMHNYLWLFLEFENFROSINSYTOAGVOMRNLGSIQPLRPGFKLFSCPSILSS	240
Dp	181	VAGTRDMHNYLWLFLEFENFROSINSYTOAGVOMRNLGSIQPLRPGFKLFSCPSILSS	240
Qy	241	WDYRRRPLANFVFLVEMGFTWBARLLISGPCDLPASASQAGITGVSHHARLLFNFC	300
Dp	241	WDYRRRPLANFVFLVEMGFTWBARLLISGPCDLPASASQAGITGVSHHARLLFNFC	300
Qy	301	LFEMESHVYTOAGVOMRNLGSIQPLRPGIKRRSCLSLPSMSWYGHLPHPHPANCFITRG	360
Dp	301	LFEMESHVYTOAGVOMRNLGSIQPLRPGIKRRSCLSLPSMSWYGHLPHPHPANCFITRG	360
Qy	361	VSPLYLSGMSQTPDLR 375	
Dp	361	VSPLYLSGMSQTPDLR 375	

RESULT 4
 PCT-US95-17111A-121
 : Sequence 121. Application PC/TUS9517111A
 :
 : GENERAL INFORMATION:
 : APPLICANT: de la Monte, Suzanne
 : APPLICANT: Wands, Jack R.
 : TITLE OF INVENTION: Neural Thread Protein Gene Expression and
 : TITLE OF INVENTION: Detection of Alzheimer's Disease
 : NUMBER OF SEQUENCES: 121
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 : STREET: 1100 New York Avenue, Suite 600
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: U.S.A.
 : ZIP: 20005-3934
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US95/17111A
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/340,426
 : FILING DATE: 14-NOV-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Ludwig, Steven R.
 : REGISTRATION NUMBER: 36,203
 : REFERENCE/DOCKET NUMBER: 0609.3840002
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 371-2600
 : TELEFAX: (202) 371-2540
 : INFORMATION FOR SEQ ID NO: 121:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 397 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : PCT-US95-17111A-121

Query Match	69.68;	Score 1415.5;	DB 5;	Length 397;
Best Local Similarity	74.68;	Pred. No. 2.8e-147;		
Matches 285;	Conservative 10;	Mismatches 44;	Indels 43;	Gaps 7

```

QY      1 MEFSLLEPRLECNCAISAHNRNLRLEGSSDSPASASPACITGMCSTHARILIFVLEMER 60
Db      1 MEFSLLEPRLECNCAISAHNRNLRLEGSSDSPASAPGVGDYRVHNAHARILIFVLEMER 60
QY      61 LHVQAGLEL-----PTSDPPVSASASARYRTGHNHARILCANFCGRNRVSLMCP 110
Db      61 LHVQAGAGLELRQMLPRSRPKVULDTGSLTMWG-----LCIANFCGRNRVSLMCP 110
QY      111 SMSPELKOSTCLSLPKCWDYRRAAVBGLIFELFLRHRCPTLTQDEYQWOCDSHLSLOPSTRE 170
Db      111 SMSPELKOSTCLSLPKCWDYRRAAVBGLIFELFLRHRCPTLTQDEYQWOCDSHLSLOPSTLR 170
QY      171 ----IKHRRPASQVAGTKRDMHNHYWMLFIFFENFLRQSLNSYVQAGVQWQWENIGSLQPLR 226
Db      171 SSILLRPP---KVAQTKRDMHNHYWMLFIFFENFLRQSLNSYVQAGVQWQWENIGSLQPLR 226
QY      227 PGFKLFSCPSLLSMDYRRPRPLANFVFLVEMGF-----TMPEAR-----LILISGPCDLPA 278
Db      227 PGFKLFSCPSLLSMDYRRPRPLANFVFLVEMGFHNHAYQVADARSJLYICLRP----- 281
QY      279 SASQASGITGVSHNARLIFNFCLEFMESHSHVYQAGVQWQWENIGSLQPLRPGKLRKFSCLSLP 338
Db      282 --PKVLGQDVTPTARPIFNFCLEFMESHSHVYQAGVQWQWENIGSLQPLRPGKLRKFSCLSLP 339
QY      339 SSMQDGHLPHPHPANFCIFIRGG 360
Db      340 SSMQDGHLPHPHPANFCIFIRGG 361

```

```

RESULT 5
US-09-605-785-573
; Sequence 573; Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqul
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C16
; CURRENT APPLICATION NUMBER: US/09/605,785
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 573
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-605-785-573

```

Query Match	12.3%	Score	250	DB	4	Length	132
Best Local Similarity	61.5%	Pred. No.	9	le-20			
Matches	48	Conservative	6	Mismatches	24	Indels	0
				Gaps			0
OY	298	NECLFEMSHSYTQGVQVNPINIGSIQIPLPGIKRKSCLSLPSSWDYGHLPHPANPCITFL	357				
Db	27	NEEFLRQSGPAAQGVQWHDLSIQLPLPHRKQPSCLSPHSWDRYAPHPILANFCSPS	86				
OY	358	RGVSPYLSGWSQTPDLR	375				

Db 87 RDGVSLCCGWSKTPGLQ 104

RESULT 6

US-09-058-489-22
; Sequence 22, Application US/09058489
; Patent No. 6103886

GENERAL INFORMATION:

APPLICANT: Whitehead Institute for Biomedical Research

APPLICANT: Lahn, Bruce

APPLICANT: Page, David

TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of

FILE REFERENCE: WH197-08pa

CURRENT FILING DATE: 1997-04-11

EARLIER FILING DATE: 1997-04-11

NUMBER OF SEQ ID NOS: 91

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 22

LENGTH: 1079

TYPE: PRT

ORGANISM: Human

US-09-058-489-22

Query Match 10.7%; Score 218; DB 3; Length 1079;

Best Local Similarity 26.4%; Pred. No. 6.3e-15;

Matches 97; Conservative 30; Mismatches 89; Indels 152; Gaps 16;

QY 25 PGSSDPASASP-----VAGITGM-CTHARLILYFFLEMEFHHGQAG--LELPDSD 74

DB 766 PSASTATPSPKSTEGKINSVSLNSPHGL-----HTVNGEGKGSOSTKVDPLAS 821

QY 75 D-----PSVASOSANRYRGHARLCLANPCGNRVSLMCPSPSPK----- 117

DB 822 HRSTSQLPMSVS-----ICPSTEVLLKACRNPGKNGL 855

QY 118 QSTCLSLKCMVDYRAAIVGLFLFLRHRCPTLNODEVOMDHSLSPPPEI----- 171

DB 856 SNSCLLDKCPRRPTSP-----YPLPKDK-----LMPPTSTYLENKR 896

QY 172 -----KHP-----PAS 177

DB 897 DAFPPPLHQPCTNPKNPVYIRGLAGALKLDLGLSTKLYEANNHWEVRYTOLLQAD 956

QY 178 AS-QVAGTK-----DMHHTWLIFFIFNFKRQSLNSVYQAGYQMRNLGSLDPLPGF 229

DB 957 ENMDPTGKKIWRCHESNSHTTIAYK--OYQASSFQESLRAGMOWCDLSLPPPGF 1013

QY 230 KLFSCPSLISMDYRRPRLANFVFLVEMGFTMFARL---ILISGPCDLPASASOSAGI 286

DB 1014 KRSHLSLPSNMNVRHLPSCPTNFCIFVETGFHNGQACLELITSG--GLASASOSAGI 1071

QY 287 TGVSHHAR 294

DB 1072 TGVSHHAR 1079

RESULT 7

US-09-227-357-193

; Sequence 193, Application US/09227357

; Patent No. 6342581

GENERAL INFORMATION:

APPLICANT: Fischer et al.

TITLE OF INVENTION: 123 Human Secreted Proteins

FILE REFERENCE: P2010P1

CURRENT FILING DATE: 1999-01-08

EARLIER FILING DATE: 1998-07-07

EARLIER APPLICATION NUMBER: 60/051,926

EARLIER FILING DATE: 1998-07-07

EARLIER APPLICATION NUMBER: 60/051,926

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 193
LENGTH: 110
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (110)
OTHER INFORMATION: Xaa equals stop translation

US-09-227-357-193

Query Match 10.3% Score 210.5; DB 4; Length 110;
Best Local Similarity 50.0%; Pred. No. 1.6e-15;
Matches 44; Conservative 6; Mismatches 15; Indels 23; Gaps 2;

OY 299 FCLP-----EMESHVQAGVQWPNLGSLOPLPGLKRFSCSLPSSWDYGH 346
Db 18 FCLVFFELRLSLSLPKREC-----NLGSLQPPPRFRQFSCSLSLNSWDYRRP 66
OY 347 PPHPANCFIFRGVSPYLSGWSQTPDL 374
Db 67 PPHLANFCVYSRGVSSCGMPCGMSRTPD 94

RESULT 8

US-09-227-357-577
Sequence 577, Application US/09227357

Patent No. 6342581

GENERAL INFORMATION:

APPLICANT: Fischer et al.

TITLE OF INVENTION: 123 Human Secreted Proteins

FILE REFERENCE: P2010P1

CURRENT FILING DATE: 1999-01-08

EARLIER APPLICATION NUMBER: PCT/US98/13684

EARLIER FILING DATE: 1998-07-07

EARLIER APPLICATION NUMBER: 60/051,926

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,793

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,925

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,929

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,803

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,732

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,931

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,932

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,916

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,930

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,918

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,920

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,733

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,795

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,919

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,928

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/055,722

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,723

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,948

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,949

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,953

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,950

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,947

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,947

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,964

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/056,360

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,684

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,984

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/055,954

EARLIER FILING DATE: 1997-08-18

EARLIER APPLICATION NUMBER: 60/058,785

EARLIER FILING DATE: 1997-09-12

EARLIER APPLICATION NUMBER: 60/058,664

EARLIER FILING DATE: 1997-09-12

EARLIER APPLICATION NUMBER: 60/058,660

EARLIER FILING DATE: 1997-09-12

EARLIER APPLICATION NUMBER: 60/058,661

EARLIER FILING DATE: 1997-09-12

NUMBER OF SEQ ID NOS: 672

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 577

LENGTH: 56

TYPE: PRT

ORGANISM: Homo sapiens

US-09-227-357-577

Query Match 9.1% Score 186; DB 4; Length 56;

Best Local Similarity 64.2%; Pred. No. 2.9e-13;

Matches 34; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

OY 323 QPLPGLKRFSCSLPSSWDYGHLPHPANCFIFRGVSPYLSGWSQTPDL 375
Db 4 EPPPRFRFRFSCSLSSWDYRRAPPHVAFICTLSRDYGLPHWPGMSQTPDL 56

RESULT 9

US-09-265-630-13

Sequence 13, Application US/09265630

Patent No. 6432914

GENERAL INFORMATION:

APPLICANT: LEVINE, BETH C

TITLE OF INVENTION: BECLIN AND USES THEREOF

FILE REFERENCE: 50902-D/JPM/EMW

CURRENT APPLICATION NUMBER: US/09/265,630

CURRENT FILING DATE: 1999-03-09

NUMBER OF SEQ ID NOS: 23

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13

LENGTH: 500

TYPE: PRT

ORGANISM: human

US-09-265-630-13

Query Match 8.3% Score 168.5; DB 4; Length 500;

Best Local Similarity 31.0%; Pred. No. 5.8e-13;

Matches 62; Conservative 18; Mismatches 63; Indels 57; Gaps 9;

OY 123 SLPCWDYRAAVPGLFLF-----FLRHRCPPLTODEVQWCHSSLOPSTPEIK 172
Db 295 SVPEWNEINAMAGQTVLLHLALANKMGLKQRYRL-----VPGNHVSYLESLTDSK 347
OY 173 HPPASASQVAGTKMDHNTWTLFIFFNPLQSLNSVQAGVQWPNLGSLOPLPGRKFLF 232
Db 348 ELPLVYS--GGIR--FFW-----DNKAQVOWHDLGLDLPPLPGGKQF 385
OY 233 SCPSSLSSWDYRR--PPRLANF--FVPL-----VEMGFVFAFLILISGPCDPA 278
Db 386 PCLSLTSLMDYRHHAPRRANDHMAVAFILDVQVQFKEVEGETGFC---LPRYMDYER 441
OY 279 SASQAGITGVSHHARLIEN 298
Db 442 GKIEDTGGSGGSYSIKTOFN 461

```
RESULT 10
US-09-227-357-285
; Sequence 285, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
```

```
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 285
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-227-357-285
```

```
Query Match
Best Local Similarity 70.7%; Score 158; DB 4; Length 122;
Matches 29; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 322 LQPLPGIKRSCSLSSNDYGHLPHPANFCIFRIGVVS 362
DB 1 MQALPGKGFSCSLSPSRWDYGCATQHPANFCIFRDRVS 41
```

```
RESULT 11
US-09-265-630-11
; Sequence 11, Application US/09265630
; Patent No. 6432814
; GENERAL INFORMATION:
; APPLICANT: LEVINE, BETH C
; TITLE OF INVENTION: BECLIN AND USES THEREOF
; FILE REFERENCE: 50902-D/JPM/EMW
; CURRENT APPLICATION NUMBER: US/09/265,630
; CURRENT FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 388
; TYPE: PRT
; ORGANISM: human
US-09-265-630-11
```

```
Query Match
Best Local Similarity 65.1%; Score 155; DB 4; Length 388;
Matches 28; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 310 TQAGVQMPNLSGQPLPGIKRSCSLSPSRWDYGHLPHPAN 352
DB 346 SKAEVQWHLGLQPPRPFKQFPCLSLSTWDYHAPRPAN 388
```

```
RESULT 12
US-09-227-357-171
; Sequence 171, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
```


ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,200A
FILING DATE: 14-OCT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,118
FILING DATE: 21-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07300/026002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5070
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-951-200A-1

Query Match
Best Local Similarity 6.3%; Score 128; DB 3; Length 48;
Matches 24; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 306 SHSVTAQGVQWPNLGSLOPLPPGLKRFSCSLPSSMDY 343
1 SLSTVAQPGVQWCHISLQPLTSLRFGQFSCSLPSTWDY 38
Db

RESULT 15
US-08-951-200A-8
Sequence 8, Application US/08951200A
Patent No. 6013495
GENERAL INFORMATION:
APPLICANT: Schwartz, Martin A.
APPLICANT: Meredith Jr., Jere E.
APPLICANT: Takada, Yoshikazu
APPLICANT: Languiño, Lucia
TITLE OF INVENTION: METHODS OF USE FOR INTEGRIN BIC CELL
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,200A
FILING DATE: 14-OCT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,118

FILING DATE: 21-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07300/026002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5070
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-951-200A-8

Query Match
Best Local Similarity 6.0%; Score 122; DB 3; Length 53;
Matches 23; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 306 SHSVTAQGVQWPNLGSLOPLPPGLKRFSCSLPSSMDY 343
1 SLSTVAQPGVQWCHISLQPLTSLRFGQFSCSLPSTWDY 43
Db

Search completed: July 17, 2003, 18:20:42
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 17, 2003, 18:07:23 ; Search time 72 Seconds
(without alignments)
694.014 Million cell updates/sec

Title: US-09-380-203-2
Perfect score: 2034
Sequence: 1 MEFSLLPRLCEKNGAISAHK.....FIRGVSPLYLSGWSQTPDLR 375

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.GeneSeq_101002:*
1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	375	21 AAB01399	Neuron-associated
2	2034	100.0	375	23 ABB81538	Neural thread prot
3	1415.5	69.6	397	17 AAR95913	Neural thread prot
4	928	45.6	381	22 AAU30235	Novel human secret
5	927.5	45.6	382	22 AAU31818	Novel human secret
6	927.5	45.6	382	22 AAU32610	Novel human secret
7	927.5	45.6	382	22 AAU32707	Novel human secret
8	849	41.7	324	22 AAU29573	Novel human secret
9	849	41.7	324	22 AAU29579	Novel human secret
10	822.5	40.4	411	22 ABB08428	Novel human diagno

11	739	36.3	317	22 AAU33200	Novel human secret
12	735	36.1	290	22 ABB21913	Novel human diagno
13	734.5	36.1	361	22 ABB68738	Novel prostate spe
14	609	29.9	257	22 AAU01294	Human polypeptide
15	590	29.0	276	22 ABB07919	Novel human diagno
16	574	28.2	185	22 AAU06014	Human polypeptide
17	569	28.0	241	22 ABB11464	Human neuronal thr
18	563	27.7	202	22 AAU30686	Novel human secret
19	561	27.6	250	22 AAU31823	Novel human secret
20	560	27.5	296	22 AAU33304	Novel human secret
21	553	27.2	175	22 AAU31782	Novel human secret
22	550.5	27.1	215	22 AAU31513	Novel human secreto
23	546	26.8	215	22 ABB07921	Novel human diagno
24	543.5	26.7	216	22 AAU32615	Novel human secreto
25	524	25.8	304	22 AAU30391	Novel human secreto
26	522.5	25.7	175	22 AAU31857	Novel human secreto
27	521.5	25.6	418	22 AAU31980	Novel human secreto
28	520	25.6	396	22 AAU30455	Novel human secreto
29	517	25.4	213	22 AAU25396	Human protein sequ
30	504.5	24.8	196	22 AAU31786	Novel human secreto
31	503	24.7	189	22 AAU31834	Novel human secreto
32	500.5	24.6	222	22 AAU31887	Novel human secreto
33	496	24.4	229	22 AAU30354	Novel human secreto
34	495.5	24.4	213	22 AAU31789	Novel human secreto
35	488.5	24.0	293	22 AAU31676	Novel human secreto
36	484.5	23.8	378	22 AAU212915	Human polypeptide
37	482	23.7	85	21 AAB28760	Peptide #5. Homo
38	475	23.4	228	22 AAU12783	Human polypeptide
39	471.5	23.2	210	22 AAU31810	Novel human secreto
40	471	23.2	286	22 AAU28297	Novel human secreto
41	464	22.8	154	22 ABB08425	Novel human diagno
42	461	22.7	293	22 AAU17168	Human bone marrow
43	457.5	22.5	264	22 ABB22078	Novel human diagno
44	457.5	22.5	264	22 ABB22078	Novel human secreto
45	447	22.0	84	21 AAB51458	Human secreted pro

ALIGNMENTS

RESULT 1	
AAB01399	
ID	AAB01399 standard; Protein; 375 AA.
XX	
AC	AAB01399;
XX	
DF	20-OCT-2000 (first entry)
DE	Neuron-associated protein.
XX	
KW	Neuron associated protein; NEUP; neurological disorder; epilepsy;
KW	ischemic cerebrovascular disease; stroke; cerebral neoplasm;
KW	Alzheimer's disease; Pick's disease; Huntington's disease;
KW	dementia; Parkinson's disease; demyelinating disease; meningitis;
KW	prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;
KW	cerebral palsy; muscular dystrophy; central nervous system; CNS;
KW	peripheral nervous system; PNS; myopathy; schizophrenia;
KW	actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;
KW	cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
KW	myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
KW	autoimmune disease; inflammation; acquired immunodeficiency syndrome;
KW	AIDS; Addison's disease; adult respiratory distress syndrome;
KW	allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;
KW	Werner syndrome; trauma; human.
OS	Homo sapiens.
PN	WO200034477-A2.
PD	15-JUN-2000.
PF	10-DEC-1999; 99WO-US30408.
XX	

```
PR 11-DEC-1998; 9805-0210083.
PR 11-DEC-1998; 9805-9123456.
PR 09-FEB-1999; 9905-0119365.
PR 16-MAR-1999; 9905-0124687.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YF, Yue H, Baughn MR, Hillman JL, Lal P, Au-young J, Yang J;
PI Lu DM, Azimzal Y;
DR WPI: 2000-423423/36.
XX
PT New human neuron-associated proteins and polynucleotides encoding them,
PT useful for diagnosis, treatment and prevention of cell proliferative
PT disorders including cancer, neuronal and neurological disorders
XX
PS Disclosure: Page 143-144; 145pp; English.
XX
CC Human neuron-associated proteins (NEUAP) can be used for
CC treating or preventing a disorder associated with decreased
CC expression or activity of NEUAP. Antagonists of NEUAP are useful for
CC treating or preventing disorder associated with increased expression
CC or activity of NEUAP. NEUAP or their fragments or derivatives are
CC useful for treating neurological disorder such as epilepsy, ischemic
CC cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's
CC disease, Pick's disease, Huntington's disease, dementia and
CC Parkinson's disease. NEUAPs are also useful for treating other
CC demyelinating diseases, bacterial and viral meningitis, prion
CC diseases including kuru, Creutzfeldt-Jakob disease, nutritional and
CC metabolic diseases of the nervous system, neurofibromatosis, other
CC developmental disorders of the central nervous system, cerebral
CC palsy, neuromuscular disorders, autonomic nervous system disorders,
CC cranial nerve disorders, spinal cord diseases, muscular dystrophy and
CC other neuromuscular disorders, peripheral nervous system disorders,
CC inherited, metabolic, endocrine, and toxic myopathies, mental
CC disorders, including mood, anxiety and schizophrenic disorders, a cell
CC proliferative disorder such as actinic keratosis, arteriosclerosis,
CC atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective
CC tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal
CC haemoglobinuria, cancers of the adrenal gland, bladder, bone,
CC bone marrow, brain, breast, cervix, and an autoimmune/inflammatory
CC disorder such as acquired immunodeficiency syndrome (AIDS), Addison's
CC disease, adult respiratory distress syndrome, allergies, ankylosing
CC spondylitis, amyloidosis, anemia, asthma, Werner syndrome,
CC complications of cancer, hemodialysis, and extracorporeal circulation,
CC viral, bacterial, fungal parasitic, protozoal, and helminthic
CC infections, and trauma. This protein was designated g3002527.
XX
SO Sequence 375 AA:
Query Match 100.0%; Score 2034; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 2,1e-198;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEFSLLEPRLECNCAISAHNRLPPSSSDPASASPVAGTGMCTHARLLIYFLEVMEEF 60
DB 1 MEFSLLEPRLECNCAISAHNRLPPSSSDPASASPVAGTGMCTHARLLIYFLEVMEEF 60
QY 61 LHWGAGLELPTSDPVSASOSARVRTGHHARLLCLANFCGRNVNLSMCSWSPKOST 120
DB 61 LHWGAGLELPTSDPVSASOSARVRTGHHARLLCLANFCGRNVNLSMCSWSPKOST 120
QY 61 LHWGAGLELPTSDPVSASOSARVRTGHHARLLCLANFCGRNVNLSMCSWSPKOST 120
DB 61 LHWGAGLELPTSDPVSASOSARVRTGHHARLLCLANFCGRNVNLSMCSWSPKOST 120
QY 121 CUSLPKCMYDRAAVPELFLFLFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPASASQ 180
DB 121 CUSLPKCMYDRAAVPELFLFLFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPASASQ 180
QY 121 CUSLPKCMYDRAAVPELFLFLFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPASASQ 180
DB 121 CUSLPKCMYDRAAVPELFLFLFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPASASQ 180
QY 181 VAGTKDMHHTYTWLFIFFNFELRQSLNSTQAGVQWRNLGSLQPLPGFKFSCPSLSS 240
DB 181 VAGTKDMHHTYTWLFIFFNFELRQSLNSTQAGVQWRNLGSLQPLPGFKFSCPSLSS 240
QY 181 VAGTKDMHHTYTWLFIFFNFELRQSLNSTQAGVQWRNLGSLQPLPGFKFSCPSLSS 240
DB 181 VAGTKDMHHTYTWLFIFFNFELRQSLNSTQAGVQWRNLGSLQPLPGFKFSCPSLSS 240
QY 241 WQYRRPRLANFVFLVEMGTFMFARLILISGCPDLPASASOSAGITGVSHARLLIFNFC 300
DB 241 WQYRRPRLANFVFLVEMGTFMFARLILISGCPDLPASASOSAGITGVSHARLLIFNFC 300
XX 241 WQYRRPRLANFVFLVEMGTFMFARLILISGCPDLPASASOSAGITGVSHARLLIFNFC 300
```

```
QY 301 LPEMESHVTOAGVQWNLGSLQPLPGFKFSCSLPPSSMDVGHLPHPANCFIRGC 360
DB 301 LPEMESHVTOAGVQWNLGSLQPLPGFKFSCSLPPSSMDVGHLPHPANCFIRGC 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375
XX
RESULT 2
ABB81538
ID ABB81538 standard; Protein: 375 AA.
XX
AC ABB81538;
XX
DT 02-SEP-2002 (first entry)
XX
DE Neural thread protein (NTP).
XX
KW Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
KW hypoxia; ischaemia; cerebral infarction.
XX
OS Homo sapiens.
XX
PN MO200234915-A2.
XX
PD 02-MAY-2002.
XX
PE 25-OCT-2001; 2001MO-US42813.
XX
PR 27-OCT-2000; 2000US-0697590.
XX
PA (NYMO-) NYMOX PHARM CORP.
PI Fitzpatrick J, Averbach P, Focht MSS, Bibiano R;
XX
DR WPI: 2002-507998/54.
DR N-FSDB; ABB89470.
XX
PT New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
PS Example 1; Fig 1; 53pp; English.
XX
CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (1).
CC (1) have neuroprotective, neurotropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases. The present sequence represents a neural thread protein
CC given in the exemplification of the present invention.
XX
SO Sequence 375 AA:
Query Match 100.0%; Score 2034; DB 23; Length 375;
Best Local Similarity 100.0%; Pred. No. 2,1e-198;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEFSLLEPRLECNCAISAHNRLPPSSSDPASASPVAGTGMCTHARLLIYFLEVMEEF 60
DB 1 MEFSLLEPRLECNCAISAHNRLPPSSSDPASASPVAGTGMCTHARLLIYFLEVMEEF 60
XX 1 MEFSLLEPRLECNCAISAHNRLPPSSSDPASASPVAGTGMCTHARLLIYFLEVMEEF 60
```

Db 1 MESSLLPRLPCNGAISAHNRRLRPGSSDSPASAPVAGITGMCTHARLLIYFLVEMEF 60
 QY 61 LHWGAGLELPTSDPVSASASQARYRTGHHARLCLANFCGRNVSLMCPSMSELKOST 120
 Db 61 LHWGAGLELPTSDPVSASASQARYRTGHHARLCLANFCGRNVSLMCPSMSELKOST 120
 QY 121 CLSLPKCMQDYRAAVPGLFLFLRHRCPRTLQDEVOMCHSSIQPSTPEIKHPASASQ 180
 Db 121 CLSLPKCMQDYRAAVPGLFLFLRHRCPRTLQDEVOMCHSSIQPSTPEIKHPASASQ 180
 QY 181 VASTKDMHNTWLIFFIFENFLRQSLNSVTQAGVQMPNLGSLQPLPPGFKLFSCPSLSS 240
 Db 181 VASTKDMHNTWLIFFIFENFLRQSLNSVTQAGVQMPNLGSLQPLPPGFKLFSCPSLSS 240
 QY 241 WDTRRRPRLANFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIFNFC 300
 Db 241 WDTRRRPRLANFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIFNFC 300
 QY 301 LFEMESHSTVQAGVQMPNLGSLQPLPPGLKRFSCSLSPSSMDYGHLPPHPANFCIFIRGG 360
 Db 301 LFEMESHSTVQAGVQMPNLGSLQPLPPGLKRFSCSLSPSSMDYGHLPPHPANFCIFIRGG 360
 QY 361 VSPYLSGMSQTPDLR 375
 Db 361 VSPYLSGMSQTPDLR 375

RESULT 3
 AAR95913 standard; Protein: 397 AA.
 ID AAR95913
 AC AAR95913;
 XX 13-NOV-1996 (first entry)
 DF 13-NOV-1996 (first entry)
 DE Neural thread protein.
 XX Neural thread protein.
 KW Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
 KM monoclonal antibody; binding fragment.
 OS Homo sapiens.
 XX Homo sapiens.
 PN WO9615272-A1.
 PD 23-MAY-1996.
 PF 14-NOV-1995; 95WO-US17111.
 XX 14-NOV-1994; 94US-0340426.
 PR 14-NOV-1994; 94US-0340426.
 XX (GEHO) GEN HOSPITAL CORP.
 PA (GEHO) GEN HOSPITAL CORP.
 PI De LA MONTE S, Wands JR;
 XX De LA MONTE S, Wands JR;
 DR WPI; 1996-259865/26.
 XX N-PSDB; AAT27738.
 PT Detection of neural thread protein in diagnosis of Alzheimer's
 PT disease - also NTP DNA and protein sequences used in gene and
 PT anti-sense therapy
 XX
 PS Claim 22; Page 171-172; 238pp; English.
 XX
 CC A method for detecting the presence of neural thread protein (NTP)
 CC having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
 CC subject comprises (a) contacting a sample from a human subject that
 CC is suspected of containing the NTP with at least one molecule
 CC capable of binding to the protein; and (b) detecting any of the
 CC molecule bound to the protein. The binding molecule is selected
 CC from an antibody free of natural impurities, a monoclonal antibody
 CC or a binding fragment of either of these. The method may be used for
 CC diagnosing the presence of Alzheimer's disease, neuroectodermal

CC tumours and a malignant astrocytoma in a human.
 XX
 SO Sequence 397 AA;
 Query Match 69.6%; Score 1415.5; DB 17; Length 397;
 Best Local Similarity 74.6%; Pred. No. 2.2e-135;
 Matches 285; Conservative 10; Mismatches 44; Indels 43; Gaps 7;
 QY 1 MESSLLPRLPCNGAISAHNRRLRPGSSDSPASAPVAGITGMCTHARLLIYFLVEMEF 60
 Db 1 MESSLLPRLPCNGAISAHNRRLRPGSSDSPASAPVAGITGMCTHARLLIYFLVEMEF 60
 QY 61 LHWGAGLELPTSDPVSASASQARYRTGHHARLCLANFCGRNVSLMCP 110
 Db 61 LHWGAGLELPTSDPVSASASQARYRTGHHARLCLANFCGRNVSLMCP 110
 QY 111 SMSPELKOSTCLSLPKCMQDYRAAVPGLFLFLRHRCPRTLQDEVOMCHSSIQPSTPE 170
 Db 111 SMSPELKOSTCLSLPKCMQDYRAAVPGLFLFLRHRCPRTLQDEVOMCHSSIQPSTLR 170
 QY 171 ----IKHPASASQVAGTKDMHNTWLIFFIFENFLRQSLNSVTQAGVQMPNLGSLQPLP 226
 Db 171 SSIILPQP-----KVAGTKDMHNTWLIFFIFENFLRQSLNSVTQAGVQMPNLGSLQPLP 226
 QY 227 PGFKLFSCPSLSSMDYRRPPRLANFVFLVEMGFTMFAR-----LILSGPCDLP 278
 Db 227 PGFKLFSCPSLSSMDYRRPPRLANFVFLVEMGFTMFAR-----LILSGPCDLP 281
 QY 279 SASQAGITGVSHHARLIFNFCLEFEMESHSTVQAGVQMPNLGSLQPLPPGLKRFSCSLP 338
 Db 282 --PKVLQDVTPYARPIFNFCLEFEMESHSTVQAGVQMPNLGSLQPLPPGLKRFSCSLP 339
 QY 339 SSMWYGHLPHPANFCIFIRGG 360
 Db 340 SSMWYGHLPHPANFCIFIRGG 361

RESULT 4
 AAU30235 standard; Protein: 381 AA.
 ID AAU30235
 AC AAU30235;
 XX 18-DEC-2001 (first entry)
 DT 18-DEC-2001 (first entry)
 DE Novel human secreted protein #726.
 XX Novel human secreted protein #726.
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 OS Homo sapiens.
 XX Homo sapiens.
 PN WO200179449-A2.
 PD 25-OCT-2001.
 PF 16-APR-2001; 2001WO-US08656.
 XX 16-APR-2001; 2001WO-US08656.
 PR 18-APR-2000; 2000US-0552929.
 XX 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX 26-JAN-2001; 2001US-0770160.
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Drmanac RT;
 XX Tang YT, Liu C, Drmanac RT;
 DR WPI; 2001-611725/70.
 XX WPI; 2001-611725/70.
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 PS Claim 20; Page 264-265; 765pp; English.
 XX

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptide and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

CC Sequence 381 AA:

Query Match 45.6%; Score 928; DB 22; Length 381;
 Best Local Similarity 55.9%; Pred. No. 8.8e-86;
 Matches 223; Conservative 27; Mismatches 101; Indels 48; Gaps 14;

QY 4 SLLPRLKCGALISAHRLNRLPGSSDSPASAPVAGITGCTHARLILYFLVEMEFYH 63
 DB 4 SLLPRLKCGALISAHRLNRLPGSSDSPASAPVAGITGCTHARLILYFLVEMGFH 62
 QY 64 GOAGLELPTSDPSVSASOSARVTRG--HNAHRLCLANFCGRNVSIMCPSM----- 112
 DB 63 AQAQLELITSDPTL-ASQSAGI-IGVNHAMLEF--FCSRDIVSLCYGWSRVAXSRIT 118
 QY 113 --SPELKOSTCLSLPKCMYRRAAV-PGLFLFLHRCPTLTODEVOMCHSSLOP-- 166
 DB 119 ATSPGLKXKACFSLPSRSDRYHVPHPHGNFCIF-----GRDEVSPC-----WPGWF 164
 QY 167 STEPIKHPASASQVAGTDMHNYT---LIFIFINFLROSLNSVTOAGVQMRNLSL 222
 DB 165 XTPLDLKRPASASQSAEIIIGVSHHTWQEVFLFLNFIYLRMSLDVSAQAARVORDLSL 224
 QY 223 QLPPEGKLFSCPSLSSWDYRR-PPRLANFVFLVEMGFMTFARLILISGPCDLPASAS 281
 DB 225 QAPPRPKFPSCSLSPSSWDYRRPPHPANFVFLVETGFTVLARVILISXPRDLPASAS 284
 QY 282 QSAGITGVSHHARLINFCLFEMESHVTOAGVQMRNLSLOPLPGLKRFSCSLSPSSW 341
 DB 285 QSAGITGVSHHARLINF--FETGTHSVTMAAVQWYITISLQPRTELKXSSHILITSNM 342
 QY 342 DYGHLPHPAN-----FCIFIRGVSPLYSGMSOTPDLR 375
 DB 343 DYRCPTPCPPNLFYLFYFHRDEGSLCCPGMSXTPELK 381

RESULT 5

AAU31818 ID AAU31818 standard; Protein; 382 AA.

AAU31818;

18-DEC-2001 (first entry)

Novel human secreted protein #2309.

Human: vaccination; gene therapy; nutritional supplement;
 stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 immune suppression; immune stimulation; anti-inflammatory; leukemia.

Homo sapiens.

WO200179449-A2.

25-OCT-2001.

XX

PF 16-APR-2001; 2001WO-US08656.
 XX 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.

XX (HSE-) HXSEQ INC.

PI Tang YF, Liu C, Drmanac RT;

DR WPI, 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 XX vaccination, testing and therapy -

PS Claim 20; Page 513; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

CC Sequence 382 AA:

Query Match 45.6%; Score 927.5; DB 22; Length 382;
 Best Local Similarity 55.8%; Pred. No. 1e-85;
 Matches 223; Conservative 27; Mismatches 101; Indels 49; Gaps 14;

QY 4 SLLPRLKCGALISAHRLNRLPGSSDSPASAPVAGITGCTHARLILYFLVEMEFYH 63
 DB 4 SLLPRLKCGALISAHRLNRLPGSSDSPASAPVAGITGCTHARLILYFLVEMGFH 62
 QY 64 GOAGLELPTSDPSVSASOSARVTRG--HNAHRLCLANFCGRNVSIMCPSM----- 113
 DB 63 AQAQLELITSDPTL-ASQSAGI-IGVNHAMLEF--FCSRDIVSLCYGWSRVAXSRIT 118
 QY 114 ----PELKOSTCLSLPKCMYRRAAV-PGLFLFLHRCPTLTODEVOMCHSSLOP-- 166
 DB 119 ATSPGLKXKACFSLPSRSDRYHVPHPHGNFCIF-----GRDEVSPC-----WPGWF 164
 QY 167 STEPIKHPASASQVAGTDMHNYT---LIFIFINFLROSLNSVTOAGVQMRNLSL 221
 DB 165 XTPLDLKRPASASQSAEIIIGVSHHTWQEVFLFLNFIYLRMSLDVSAQAARVORDLSL 224
 QY 222 LQPLPEGKLFSCPSLSSWDYRR-PPRLANFVFLVEMGFMTFARLILISGPCDLPASA 280
 DB 225 QAPPRPKFPSCSLSPSSWDYRRPPHPANFVFLVETGFTVLARVILISXPRDLPASA 284
 QY 281 SOSAGITGVSHHARLINFCLFEMESHVTOAGVQMRNLSLOPLPGLKRFSCSLSPSS 340
 DB 285 SOSAGITGVSHHARLINF--FETGTHSVTMAAVQWYITISLQPRTELKXSSHILITSN 342
 QY 341 WDYGHLPHPAN-----FCIFIRGVSPLYSGMSOTPDLR 375
 DB 343 WDYRCPTPCPPNLFYLFYFHRDEGSLCCPGMSXTPELK 382

RESULT 6

AAU32610 ID AAU32610 standard; Protein; 382 AA.

XX

```
AC AAU32610;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3101.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KM stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
XX 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 637; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration.
CC Immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 382 AA:
Query Match 45.6%; Score 927.5; DB 22; Length 382;
Best Local Similarity 55.8%; Pred. No. 1e-85;
Matches 223; Conservative 27; Mismatches 101; Indels 49; Gaps 14;
QY 4 SLLPRLCNGAISHRNRLRPGSSDPASAPVAGITGMCTHARLLIYFLVEMEPILHV 63
DB 4 SLLPRLCNGAISHRNRLRPGSSDPASAPVAGITGMCTHARLLIYFLVEMEPILHV 62
QY 64 GAGLELPDSDPSVASASARVYRTG--HNAFLCLANFCGRNRVSLMCPNS----- 113
DB 63 AAGAGLELTSDEPTL-ASQSAGI-TGVNHNHMLFF--FCSRPTVSLCYRGWRVAXSRTT 118
QY 114 -----PELKOSTCLSLPKCMQYRRAAV-PGLFTLFLRHRCPPTLTODEVQWCDHSSLQP-- 166
DB 119 ATSAGPLKXFAFCFSILPSSRDYRHRVPHGNCIF-----GRDEVSPC-----WPGW 164
QY 167 -STPIKHPPASASQVAGTKDMHNTW-----LIFIFINELPQSLNSTVQAGVQRNIGS 221
DB 165 FXTPLKRPASASQSAELTIGVSHHTWPOEYVFLPNTLYLRLSDVAQAQVRRDIGS 224
QY 222 LQPLPPGFKLSPCLSSMDYRR-PRIANFVFLVEMGFTMFARLLISGPCDLPA 280
```

```
DB 225 LQAPPREFKPSCLSLPSSMDYRRPPHPANFVFLVETGFTVLARVLISXPRDLPA 284
QY 281 SOSAGITGVSHHARILNFCLFEMESHVYTOAGVQMPNLGSLQPLPGLKRRSCSLDPS 340
DB 285 SOSAGITGVSHHTRLIFNF--FETGTHSVTAAVQWYTIIGSLQPTPELKSXSHLILTSN 342
QY 341 WDYGHLPHPAN-----FCIFIRGGSVPYLSGWSOTPDLR 375
DB 343 WDYRCTPPCPNPLFIYLFYFHRDEGSLCCPGWSXTPELK 382
RESULT 7
AAU32707
ID AAU32707 standard; Protein; 382 AA.
XX
AC AAU32707;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3198.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KM stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
XX 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 652; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration.
CC Immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 382 AA:
Query Match 45.6%; Score 927.5; DB 22; Length 382;
Best Local Similarity 55.8%; Pred. No. 1e-85;
Matches 223; Conservative 27; Mismatches 101; Indels 49; Gaps 14;
QY 4 SLLPRLCNGAISHRNRLRPGSSDPASAPVAGITGMCTHARLLIYFLVEMEPILHV 63
```


DR WPI: 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
XX
PS Claim 20: Page 168: 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation: to regulate haematopoiesis; and in
CC bone, cartilage, tendon, and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AA029510-AA03304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 324 AA:
Query Match 41.7%; Score 849; DB 22; Length 324;
Best Local Similarity 50.8%; Pred. No. 7.8e-78;
Matches 200; Conservative 10; Mismatches 62; Indels 122; Gaps 10;
QY 1 MEFSLLPRLECGAISAHRNRLPGSSDPSASPAGITGCTHARLLYFLVEMEF 60
DB 34 MEFSLLPRLECGSRTISAQVNLRLPCSSDSSASASVAGITGCHHAQLI-FVFLVEKGF 92
QY 61 LHVGAQGLELPTSDPSVSASOSARYRIGHARILCLANFCGRNRVSLMCPSPSEIKOST 120
DB 93 HHVGAQGLELTATASGDP-----T 109
QY 121 CLSLKCMQDYRR-AAVPGLEFLFRLHRCPTLTODEVQCMCHSLQSPSTPEIKHPPASAS 179
DB 110 CGLRPMQDYRRKXATAPGLE----- 129
QY 180 QVAGTKDMHHYTWLI-FIFLNFRLQSLNSVTQAGVQWRNLGSLQ-PLPPGKFLSCPSLL 238
DB 130 -----FFFLROSFTLVQAQAGVQWRDLGSLQPPRRFRQFSLSLP 170
QY 239 S---SMDYRR-PPRLAN-----FVFLVEMGFTMFARLLISGP-CDLPASASOSAGIT 287
DB 171 STMSVYRRHAPPCPANFVFFFEFFFLVEKGFSSMLRLTLVNLXPHDPPDPASASOSAGIT 230
QY 288 GSHHAR-----LIFNCLFEMESHVSVTQAGVQWRNLGSLQELPGLKRFSCLSIPSSW 341
DB 231 GVSHHTRPMSFNKNTIFFEFTETESRVAQAGVQWRDLSSRQPPGKFRFSCLSLSSW 290
QY 342 DYGHLLRPPRANCFIRGCVSPYLSGWSQTPDLR 375
DB 291 DYRVRPMPMGXCFIRROGCVSPRMSGWSQTPDLK 324
RESULT 10
ABG08428
ID ABG08428 standard; Protein; 411 AA.
XX
XX ABG08428;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #8419.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX

OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) 'HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR N-PSDB: AAS72615.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
PS Claim 20: SEQ ID NO 38787; 103pp; English.
XX
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 411 AA:
Query Match 40.4%; Score 822.5; DB 22; Length 411;
Best Local Similarity 51.0%; Pred. No. 5.4e-75;
Matches 200; Conservative 37; Mismatches 120; Indels 35; Gaps 14;
QY 4 SLLPRLECGAISAHRNRLPGSSDPSASPAGITGCTHARLLYFLVEMEF 63
DB 12 SPLRLQCSGMLAYCNRLLGSSNSVASAGVAGITGCHHTOXI-FVFLVEYGFHHV 70
QY 64 GQAGLELPTSDPSVSASOSAR-YRIGHAR--LCLANFCGRNRVSLMCPSPSEIKOST 120
DB 71 GQAGLELPTSGDLPYSASOSARDYRDDITREKC--FCXKTHILMSPSSSEERSQ 127
QY 121 CLS-----LPKCMQDYRRRAVPGLEFLFRLHRCPTLTODEVQCMCHSLQSPSTPEIK 172
DB 128 FFSQDSXKSSRYPK-----KLGMVYIAINITYGRYTKKXKFKREYIGH-SIHRN 179
QY 173 HPPASASOYAGTKDMHHYTWLI--FIFLNFRLQSLNSVTQAGVQWRNLGSLQPLPGF 229
DB 180 H-GVNTYHVAASEK-LHXYYTRMKKAKNFFFLRLRELASVAQAGVQWRDLGSLQSPPGF 237
QY 230 KLFSCPSLLSMDYRR--PPRLANFVFLVEM--GFTMFARLLISGPDLPASASOSAG 285
DB 238 TPFSCPASLRSDYRRVRLPCLANFLLYFSMRGRFTYFSRMVYSISKPDOPASASOSAG 297

QY 286 ITGVSHART--IFNFCLF-EMESHVTOAGVOMPNLSLOPLPGLKRFSCSLSPSSMD 342
DB 298 ITGVSHARKKIIFFFEFSETESRSVAQGVOMHDLGSRHRPPGFTFFSCLSFPSSMD 357
QY 343 YGHLPPHPANPCIFIRGVSPLYLSGNSQTPDL 374
DB 358 YRGPPRPANPCVFSRSDGSPCKGWSRSPDL 389
RESULT 11
AAU33200 ID AAU33200 standard; Protein; 317 AA.
XX
AC AAU33200;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3691.
XX
KW Human: vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PE 16-APR-2001; 2001MO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
PS Claim 20; Page 729; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon, and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 317 AA:
Query Match 36.3%; Score 739; DB 22; Length 317;
Best Local Similarity 54.8%; Pred. No. 1,2e-66;
Matches 172; Conservative 17; Mismatches 93; Indels 32; Gaps 8;
QY 1 MEFSLPRLCEGSAISAHNRLLPGSSDPASAPVAGITGCTHARLITFFLVEMEF 60
DB 19 MEFSLPRLCEGSAISAHNRLLPGSSDPASAPV-GMDYRHYVHPRANVFEEFSMDG 77

QY 61 LHVQAGU-----ELPTSDPSVASOSARYRIGHARLCLANPCGRNRVSLMCPMS 113
DB 78 VSPCHSGLYNSRRQMTIPSRPRKVLDTGLATM-----PGCLLANRCGRNRVSLMCPMS 132
QY 114 PELKOSTCLSLPKCWDYDRAAVPGFLFELHNRCPITLQDEVOMCHSLOPSTEIKH 173
DB 133 PELKOSTCLSLPKCWDYDRAAVPGFLFELHNRCPITLQDEVOMCHSLOPSTEIKH 192
QY 174 PPASASQVAGTKDMHAYTWLI--FIFENFLRQSLNSVTOAGVOMNLSLOPLPFGFRL 231
DB 193 PPASASQ--SSMDORHAPLHLANFEYFEXFETESHVTRLECSGALIANCLLPGSSY 250
QY 232 FSCPSLLSSMDYRRPRLANFEVFLVEMGF---TMFAR---LILSGPCLDPPASASOS 283
DB 251 SPAPASVXVAGTTGAHRRILANFEVFLVEMGFHVRQVDARSIDLVLCLPRP-----PKV 303
QY 284 AGITGVSHARLIF 297
DB 304 LGLDVSHHRRAYF 317
RESULT 12
ABG21913 ID ABG21913 standard; Protein; 290 AA.
XX
AC ABG21913;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21904.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001MO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS86100.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 52272; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.
Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 290 AA:

Query Match 36.1%; Score 735; DB 22; Length 290;
Best Local Similarity 47.6%; Pred. No. 2.7e-66;

Matches 182; Conservative 17; Mismatches 69; Indels 114; Gaps 12;

```
OY 6 LLPRLECGAISAHNRLPGSSDSPASAPVAGITGCTHARLITFLVEMERLHVGQ 65
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 11 LSPRLCEGSAISAHNRLCPSSSDSPASR--ITGISGARHHAQLILVFLVEMRHVHGQ 68
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 66 AGLEPTSDPSVSASOSARYRTGHHARLCLANFCGRNRVSLMCPSPMELKOSTCLSLP 125
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 69 AGLKLTSGNP-----HH-----LGLP 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 126 KCMQYRAAVGGLFLFLHRHRCPLTODEVQWCDHSLQSTPEIKHPASASOVAGTK 185
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 86 KCMQYSRKPP-----RPPRP----- 101
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 186 DMHHTWLIFIFIFLRLQSLNVTQAGVQWNRIGSLQPLPPGFRLFCSPSLSSMDYR- 244
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 102 -----ITFFFLKWSLALVAQAGWQDDLSLQPLPPGFRLFCSPSLSSMDYR- 152
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 245 RPPRLANF-----VFLVEMGFTMEFARL--ILISGPCDLPA-SASOSAGITGVSHHARL 295
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 153 PPRRLANVVCVCLVVDGFTMLARGLDELITSG--DLPPFGLPKCLGFTGMCHCARP 210
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 296 IFNCLFEMESHSTQAGVQWPNIGSLQPLPPGLKRFSCSLSPSSMDYGHLPHPANFC- 354
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 211 IFFF--FEMESCFAVQAGVQWCDLSLQPLPPRFKXFCSLSLSSMDYRHMPCLANCF 268
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 355 -IFIRGVSPLYSGWSOTPDRL 375
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 269 GTFSDRVSCGMPGRSOTPDRL 290
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 13
ABG68738 standard: Protein: 361 AA.

AC ABG68738;

DT 07-OCT-2002 (first entry)

DE Human prostate specific protein DEX0293_107.

Human; prostate specific nucleic acid; PSNA; prostate cancer; PSP;
prostate specific protein; cytosolic; non-cancerous prostate disease;
gene therapy; cancer; immunostimulant; vaccine.

OS Homo sapiens.

FN W020025735-A2.

PD 18-JUL-2002.

PF 27-NOV-2001; 2001WO-USA4363.

PR 27-NOV-2000; 2000US-253176P.

PA (DIAD-) DIADEXUS INC.

PI Salceda S, Machina RA, Recipon H, Cafferkey R, Ali S, Sun Y;

DR WPI; 2002-557831/59.

DR N-PSDB; ABK97631.

PT New prostate specific genes, useful for treating or diagnosing cancer,
or useful as vaccines for treating cancer, particularly prostate
cancer, in a patient

Claim 11; Page 207-208; 212pp; English.

The invention relates to a new isolated prostate-specific nucleic acid (PSNA) molecule comprising the cDNA sequences appearing as ABK97574-ABK97642 which encode prostate specific proteins appearing as ABG68701-ABG68746, or a sequence hybridizing to a PSNA or which has 60% sequence homology with a PSNA. Also included are a method of determining the presence of a PSNA in a sample, a vector comprising the PSNA, a host cell comprising the vector, producing the polypeptide encoded by the PSNA, a method of determining the presence of a PSP in a sample, diagnosing and monitoring the presence and metastases of prostate cancer in a patient, a kit for detecting a risk of cancer or presence of cancer in a patient (the kit comprising a means for determining the presence of the PSNA or PSP in a sample of a patient) and a vaccine comprising the polypeptide or the nucleic acid encoding the polypeptide. The PSNA, PSP and anti-PSP antibody are useful for diagnosing and treating cancer in a patient (e.g. by gene therapy). The nucleic acid molecule and polypeptide are also useful as vaccines for treating cancer, particularly prostate cancer and non-cancerous prostate diseases. The present sequence is a PSP of the invention.

Sequence 361 AA:

Query Match 36.1%; Score 734.5; DB 23; Length 361;
Best Local Similarity 43.1%; Pred. No. 4.1e-66;

Matches 185; Conservative 24; Mismatches 93; Indels 127; Gaps 12;

```
OY 6 LLPRLECGAISAHNRLPGSSDSPASAPVAGITG-----CTHARL----- 49
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 LLPRLECGITMAHOKRLKILGSGDLPASASRVGGTGMROPTMGTSCHAFQIGLFPVAN 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 50 -----IIFYFLVEMERLHVGQAGLELTPSDPSVSASOSARYRTGHHARLCL 97
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 FERSFYILGTSPBLDLMLINMF-----FALLIHVLF- 94
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 98 NFCCGRNRVSLMCP--SMSPELKOSTCLSLPKCMQYRAAVPGLFLFLHRHC----- 148
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 95 -----NRDSTCCPGASLTGLFAFACSLPRCMQDYTPRL-LPIDIFLQCCLSFYPLN 148
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 149 PTLTODEVQWCDHSLQSTPEIKHPASASOVAGTKDMHHTWLIFIFINFLQSLNS 208
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 149 TVFQRAEV-----LIFQWMPVYLKWSLHS 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 209 VTOAGVQWNRIGSLQPLPPGFRLFCSPSLSSMDYR--PPRLAN--FVFLVEMGFTMEFA 265
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 173 VAQAGVQWNCNIGSLQPLPPRRRRRFSCLSLSSMDHRAHAPCLANFLFKFLVDOSFTWLA 232
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 266 RLILISGPC-DLPASASOSAGITGVSH-----HARLIENFCLFEMES 306
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 223 RLVLMSARSGLDLPAPASOSAGITGVHCTWPKSFADSHIGLAFHFAFFFFFAVAYS 292
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 307 HSYVQAGVQWRNLGSLQPLPPGLKRFSCSLSPSSMDYGHLPHPANFCITFIRGVSPILS 366
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 293 HPVQAGVQWNRDLGSLQPLPPPGFKOFLCLSLPGSDYRRAPRQANFCIFSDGVSPCWT 352
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 367 GMSOTPDRL 375
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 353 GMSOTPDRL 361
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 14
AAO10294

ID AAO10294 standard: Protein: 257 AA.

AC AAO10294;

DT 06-NOV-2001 (first entry)

Db	66	LKLPYSGDLPASASOSAGI-TGVSHRAR-----	92
Qy	126	KCMDYRAAAYPGFLFLFFLRHRCPTLTODEVOMCDHSSLOPSTPEIKHPPASASQVAGTK	185
Db	93	-----PASSYLFFF-----XDEVSLC-----	108
Qy	186	DMHHYTWLIFLIFENFLQSLNSTOAGVOMBNLGSLOPLPPGFKLFSCPSLSSMDYRR	245
Db	109	-----CQAGVQLHDVGSLO-LPPRFKRFSCPASRSSXDERR	143
Qy	246	-PPRLANFVFLVENGFTMFARLILISGP-CDLPASASOSAGITGVSHHARLIFNFCLE	303
Db	144	LPRLPNFCIFMXRMGFTILARLVNSRPOCDLPASASOSAGITGVSHHARLPASSYCIFF	203
Qy	304	MESH-SYTOAGVOMBNLGSLOPLPPGFKRFSCSLPSSMDYGHLPHPANFCIFIRGYS	362
Db	204	LHMLTIAQAGVXXCNCSSLOPPTPGFKWVSCSLPSSMDYRSLPPLPANFCIFSRDYS	263
Qy	363	PYLSGMSQTPDLR	375
Db	264	LCSLMSRTPDFK	276

Search completed: July 17, 2003, 18:17:22
Job time : 74 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2003, 07:50:41 ; Search time 338 Seconds

(without alignments)
8801.323 Million cell updates/sec

Title: US-09-380-203-1

Perfect score: 1442

Sequence: 1 TTTTGTGATGAGTGCAG.....TTAACAAGCTTAGACCA 1442

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_MA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PC7_NEM_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PC7_NEM_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1442	100.0	1442	11	US-09-964-666-1
2	1442	100.0	1442	11	US-09-964-412-1
3	1442	100.0	1442	12	US-09-964-667-1
4	1442	100.0	1442	12	US-09-872-968-1
5	1442	100.0	1442	15	US-10-146-130-1
6	1442	100.0	1442	15	US-10-092-934-1
7	1442	100.0	1442	15	US-10-153-334-53
8	1442	100.0	1442	15	US-10-198-069-48
9	1442	100.0	1442	15	US-10-198-070-125
10	1223.4	84.8	1418	11	US-09-964-666-4
11	1223.4	84.8	1418	11	US-09-964-412-4
12	1223.4	84.8	1418	12	US-09-964-667-4
13	1080.2	74.9	1381	11	US-09-964-666-3
14	1080.2	74.9	1381	11	US-09-964-412-3
15	1080.2	74.9	1381	12	US-09-964-667-3
16	587.8	40.8	65608	10	US-09-962-436-292

17	587.8	40.8	65608	11	US-09-962-832-119	Sequence 119, App
18	587.8	40.8	65608	11	US-09-954-531-180	Sequence 180, App
19	524.4	36.4	62944	11	US-09-954-456-2257	Sequence 2257, App
20	513	35.6	21470	11	US-09-764-847-1157	Sequence 1157, App
21	513	35.6	21470	15	US-10-092-154-1157	Sequence 1157, App
22	490.4	34.0	99014	11	US-09-880-107-3428	Sequence 3428, App
23	479.2	33.2	51719	10	US-09-918-686-2	Sequence 2, Appl1
24	479.2	33.2	92139	10	US-09-918-686-1	Sequence 1, Appl1
25	477.4	33.1	11319	12	US-09-764-872-579	Sequence 579, App
26	475.4	33.0	14796	11	US-09-954-456-973	Sequence 973, App
27	475.4	33.0	14796	11	US-09-954-456-1636	Sequence 1636, App
28	475.4	33.0	14796	11	US-09-918-186A-3	Sequence 3, Appl1
29	475.4	33.0	14796	11	US-09-880-107-3421	Sequence 3421, App
30	475.4	33.0	14796	15	US-10-138-618-35	Sequence 35, Appl1
31	475.4	33.0	26657	10	US-09-810-673A-3	Sequence 3, Appl1
32	471.2	32.7	12542	11	US-09-764-864-1774	Sequence 1774, App
33	467	32.4	15041	10	US-09-764-869-1421	Sequence 1421, App
34	467	32.4	15041	15	US-10-091-504-1421	Sequence 1421, App
35	464.8	32.2	57130	11	US-09-835-081-3	Sequence 3, Appl1
36	447.8	31.1	14176	11	US-09-764-864-1644	Sequence 1644, App
37	440.8	30.6	13224	10	US-09-764-853-897	Sequence 897, App
38	440.6	30.6	23106	12	US-09-863-049A-1	Sequence 1, Appl1
39	440.2	30.5	110096	11	US-09-880-107-1542	Sequence 1542, App
40	439.2	30.5	110096	11	US-09-880-107-1542	Sequence 1542, App
41	436.4	30.3	118067	14	US-10-081-327-32	Sequence 32, Appl1
42	436.2	30.2	2226	15	US-10-027-632-102037	Sequence 102037, App
43	433.2	30.0	20247	11	US-09-764-877-2680	Sequence 2680, App
44	433	30.0	32148	12	US-09-764-891-6806	Sequence 6806, App
45	429.8	29.8	27148	10	US-09-764-860-1046	Sequence 1046, App

ALIGNMENTS

RESULT 1

US-09-964-666-1

Sequence 1, Application US/09964666

Patent No. US20020104108A1

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

Wands, Jack R.

TITLE OF INVENTION: Transgenic Animals and Cell Lines for Screening Drugs Effective for the Treatment or Prevention of Alzheimer's Disease

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.

STREET: 1100 New York Ave., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/964,666

FILING DATE: 28-Sep-2001

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0609.4370000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1442 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 52,893
REFERENCE/DOCKET NUMBER: 0609,4370000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1442 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 15..1139
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-964-667-1

Query Match 100.0%; Score 1442; DB 12; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTTTGTGAGTGGAGTTTTCCTGTTGTCCTGAGCTGAGTGCATGCGCGCA 60
DB 1 TTTTGTGAGTGGAGTTTTCCTGTTGTCCTGAGCTGAGTGCATGCGCGCA 60
QY 61 TCTAGCTCACCCCACTCCGCTCCGGGTTCAAGCATTCCTGCTCAGCTCC 120
DB 61 TCTAGCTCACCCCACTCCGCTCCGGGTTCAAGCATTCCTGCTCAGCTCC 120
QY 121 CAGTGGTGGATTAGAGCTGACCCAGCTGGCTGATTTGTTGTTTATTTT 180
DB 121 CAGTGGTGGATTAGAGCTGACCCAGCTGGCTGATTTGTTGTTTATTTT 180
QY 181 TAGAGTGAATTCCTCATGTTGTCAGCTGCTCTGCACTCCGACCTAGAT 240
DB 181 TAGAGTGAATTCCTCATGTTGTCAGCTGCTCTGCACTCCGACCTAGAT 240
QY 241 CCTCGCTCGGCTCCCAAGTGTCTAGATGACAGTGGCCACATGCGGCTG 300
DB 241 CCTCGCTCGGCTCCCAAGTGTCTAGATGACAGTGGCCACATGCGGCTG 300
QY 301 TGGCTAATTTTGTGTAGAAACAGAGTTTCTGATGTCGCCAAGCTGTC 360
DB 301 TGGCTAATTTTGTGTAGAAACAGAGTTTCTGATGTCGCCAAGCTGTC 360
QY 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGAGTGCAGGCTGAG 420
DB 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGAGTGCAGGCTGAG 420
QY 421 CTGGCTTTTATTTTATTTTATTTTAAACACAGGTCCTTACCCAGATG 480
DB 421 CTGGCTTTTATTTTATTTTATTTTAAACACAGGTCCTTACCCAGATG 480
QY 481 TGCAGTGTGTGATCAGCTCAGTGCAGGCTTCACTCTGAGATCAAGCAT 540
DB 481 TGCAGTGTGTGATCAGCTCAGTGCAGGCTTCACTCTGAGATCAAGCAT 540
QY 541 CCTAGGCTCCCAAGTGTGAGGACCAAGATGACCACTGACCTGGTAA 600
DB 541 CCTAGGCTCCCAAGTGTGAGGACCAAGATGACCACTGACCTGGTAA 600
QY 601 TTTTATTTTATTTTATTTTATTTTAAACAGAGTCTGACCCAGGCTG 660
DB 601 TTTTATTTTATTTTATTTTATTTTAAACAGAGTCTGACCCAGGCTG 660
QY 661 GGGCAATCTTGGCTACTGCAACTGCTCCCGGGTCAAGTATTTCTCT 720
DB 661 GGGCAATCTTGGCTACTGCAACTGCTCCCGGGTCAAGTATTTCTCT 720
QY 721 GCTCTCTAGTGGAGTACAGGCGCCACGACTAGTATTTTGTATTTT 780

DB 721 GCTCTCTAGTGGAGTACAGGCGCCACGACTAGTATTTTGTATTTT 780
QY 781 TACTAGATGGAGTTTCAACATGTTGCGCAGTTGATCTGACCTTGTATC 840
DB 781 TACTAGATGGAGTTTCAACATGTTGCGCAGTTGATCTGACCTTGTATC 840
QY 841 TGGCTCCGCGCTCCCAAGTGTGAGTATTAAGGCTGAGCAGCAGCCGCT 900
DB 841 TGGCTCCGCGCTCCCAAGTGTGAGTATTAAGGCTGAGCAGCAGCCGCT 900
QY 901 TTTTATTTTATTTTATTTTATTTTAAATGGAATTCACCTGTTACCA 960
DB 901 TTTTATTTTATTTTATTTTATTTTAAATGGAATTCACCTGTTACCA 960
QY 961 GGGCAATCTGCGCTCACTGCACTGCACTGCTCCGCTGAGGCTGAG 1020
DB 961 GGGCAATCTGCGCTCACTGCACTGCACTGCTCCGCTGAGGCTGAG 1020
QY 1021 GCTCCCAAGCAGCTGGGATTAAGGCTGCACTGCACTGCACTGCACT 1080
DB 1021 GCTCCCAAGCAGCTGGGATTAAGGCTGCACTGCACTGCACTGCACT 1080
QY 1081 TCAATTAGAGGCGGCTTCAACATATTTGTCAGGCTGCTCAAACTCT 1140
DB 1081 TCAATTAGAGGCGGCTTCAACATATTTGTCAGGCTGCTCAAACTCT 1140
QY 1141 GACCCACCTGCTCAGCTTCAAAAGTGTGAGGATTAAGGCTGAGCAG 1200
DB 1141 GACCCACCTGCTCAGCTTCAAAAGTGTGAGGATTAAGGCTGAGCAG 1200
QY 1201 CGGCTAATTTATTAATAAATATGATGAGGAGGCTGCTGATTTT 1260
DB 1201 CGGCTAATTTATTAATAAATATGATGAGGAGGCTGCTGATTTT 1260
QY 1261 GGTCTCAAACTTGTGCTTCAATGCAATCTTCCAAATGAGCCACCA 1320
DB 1261 GGTCTCAAACTTGTGCTTCAATGCAATCTTCCAAATGAGCCACCA 1320
QY 1321 CATTTTAAACAGTTTACATTTTATTTATTTATTTATTTATTTATTT 1380
DB 1321 CATTTTAAACAGTTTACATTTTATTTATTTATTTATTTATTTATTT 1380
QY 1381 CAACCTGCAATTTAGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 CAACCTGCAATTTAGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 CA 1442
DB 1441 CA 1442

RESULT 4
US-09-872-968-1
; Sequence 1, Application US/09872968
; Publication No. US20030050262A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/09/872,968
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-872-968-1
Query Match 100.0%; Score 1442; DB 12; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1442: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTTTGTGAGATGAGTTTTCGCTGTTGCCAGGCTGAGTGAATGGCGAA 60
Db 1 TTTTGTGAGATGAGTTTTCGCTGTTGCCAGGCTGAGTGAATGGCGAA 60

Oy 61 TCTCAGCTACACCGAACCTCCGCGCTCCGGGTTCAAGGATTCCTGCTCAGCCTCC 120
Db 61 TCTCAGCTACACCGAACCTCCGCGCTCCGGGTTCAAGGATTCCTGCTCAGCCTCC 120

Oy 121 CAGTAGCTGGGATTAACAGGATGTGCACCCAGCTCCGCTAATTTTGTATTTT 180
Db 121 CAGTAGCTGGGATTAACAGGATGTGCACCCAGCTCCGCTAATTTTGTATTTT 180

Oy 181 TAGAGATGAGTTTCTCATGTTGGTGCAGGCTGGCTGGAAGTCCGACCTCAGATGATC 240
Db 181 TAGAGATGAGTTTCTCATGTTGGTGCAGGCTGGCTGGAAGTCCGACCTCAGATGATC 240

Oy 241 CCTCCGCTCGGCTCCCAAGTGTCTAGATACAGAGCTGGCCACCATGCCGGCTCTGCTC 300
Db 241 CCTCCGCTCGGCTCCCAAGTGTCTAGATACAGAGCTGGCCACCATGCCGGCTCTGCTC 300

Oy 301 TGGCTAATTTTGTGTGTAAGAACAGGTTTCTGATGTGCCAAAGCTGTCTCTGAGC 360
Db 301 TGGCTAATTTTGTGTGTAAGAACAGGTTTCTGATGTGCCAAAGCTGTCTCTGAGC 360

Oy 361 TCAAGCAGTCCAGCTGCTGAGCTCCCAAGTGTGGATTTACAGGGCTGACGCTGCTG 420
Db 361 TCAAGCAGTCCAGCTGCTGAGCTCCCAAGTGTGGATTTACAGGGCTGACGCTGCTG 420

Oy 421 CTGGCCCTTTTATTTTATTTTATTTTAAAGACAGAGTGTCCACTCTTTACCCAGATGAG 480
Db 421 CTGGCCCTTTTATTTTATTTTATTTTAAAGACAGAGTGTCCACTCTTTACCCAGATGAG 480

Oy 481 TGCAGTGTGTGATCAGAGCTCAGCTGCGCTTCACTGCTGAGATCAAGATCTCTG 540
Db 481 TGCAGTGTGTGATCAGAGCTCAGCTGCGCTTCACTGCTGAGATCAAGATCTCTG 540

Oy 541 CCTCAGCTCCCAAGTACCTGGAGCAAGACATGCACATGACCTGAGGATTTT 600
Db 541 CCTCAGCTCCCAAGTACCTGGAGCAAGACATGCACATGACCTGAGGATTTT 600

Oy 601 TTTTATTTTATTTTATTTTATTTTGAAGACAGAGTGTCACTGTGACCCAGGCTGAGTGCAGT 660
Db 601 TTTTATTTTATTTTATTTTATTTTGAAGACAGAGTGTCACTGTGACCCAGGCTGAGTGCAGT 660

Oy 661 GGGCCAAATCTGGGCTCAGTGAACCTTCCGCTCCGGGTTCAAGTATTTCTGCCCCA 720
Db 661 GGGCCAAATCTGGGCTCAGTGAACCTTCCGCTCCGGGTTCAAGTATTTCTGCCCCA 720

Oy 721 GCTCTCTGAGTACGCTGAGGCTACAGGCTCCAGCAGCTGATTTTGTATTTT 780
Db 721 GCTCTCTGAGTACGCTGAGGCTACAGGCTCCAGCAGCTGATTTTGTATTTT 780

Oy 781 TAGTAGAGATGGGGTTACACATGTTCCGACAGTTGATCTTCTGACCTGTGATC 840
Db 781 TAGTAGAGATGGGGTTACACATGTTCCGACAGTTGATCTTCTGACCTGTGATC 840

Oy 841 TGGCTGCTCGGCTCCCAAGTGTCTGGGATTAAGGGGTGAGCCACACCCGCGCTTA 900
Db 841 TGGCTGCTCGGCTCCCAAGTGTCTGGGATTAAGGGGTGAGCCACACCCGCGCTTA 900

Oy 901 TTTTAAATTTTGTGTTGTAAGTGAATCTCACTGTTTACCAGGCTGGAGTGCAT 960
Db 901 TTTTAAATTTTGTGTTGTAAGTGAATCTCACTGTTTACCAGGCTGGAGTGCAT 960

Oy 961 GGGCAATCTGGGCTCAGTGAACCTTCCGCTCCGGGTTCAAGGATTTCTGCTCA 1020
Db 961 GGGCAATCTGGGCTCAGTGAACCTTCCGCTCCGGGTTCAAGGATTTCTGCTCA 1020

Oy 1021 GCTCTCCAGAGAGCTGGGATTAAGGGGCACTGCCACACACCCGCTAATTTTGTATTT 1080
Db 1021 GCTCTCCAGAGAGCTGGGATTAAGGGGCACTGCCACACACCCGCTAATTTTGTATTT 1080

Oy 1081 TCATTAGAGGGGGTTTACCATTTTGTGAGGCTGCTCAAACTCCTGACCTCAGGT 1140
Db 1081 TCATTAGAGGGGGTTTACCATTTTGTGAGGCTGCTCCTCAAACTCCTGACCTCAGGT 1140

Oy 1141 GACCCACCTGCTCAGCTCCCAAGTGTGGGATTTACAGGCTGAGCCACCTCAGCCAG 1200
Db 1141 GACCCACCTGCTCAGCTCCCAAGTGTGGGATTTACAGGCTGAGCCACCTCAGCCAG 1200

Oy 1201 CCGGCTAATTTAGATTAAGTAAATATGTAGCAATGGGGGCTCTGCTATGTGCCAGCT 1260
Db 1201 CCGGCTAATTTAGATTAAGTAAATATGTAGCAATGGGGGCTCTGCTATGTGCCAGCT 1260

Oy 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCCAAATGAGCCACACCCAGCAGTCA 1320
Db 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCCAAATGAGCCACACCCAGCAGTCA 1320

Oy 1321 CATTTTAAACAGTTTACATCTTTATTTTATTTTACTATGAAAGTAAATTAACATGT 1380
Db 1321 CATTTTAAACAGTTTACATCTTTATTTTATTTTACTATGAAAGTAAATTAACATGT 1380

Oy 1381 CAACCTGCAATTTAGTAGTAAAGAGTCTTTTATTAACCTTTTAAACAAAGCTTTAAG 1440
Db 1381 CAACCTGCAATTTAGTAGTAAAGAGTCTTTTATTAACCTTTTAAACAAAGCTTTAAG 1440

Oy 1441 CA 1442
Db 1441 CA 1442

RESULT 5
US-10-146-130-1
; Sequence 1, Application us/10146130
; Publication No. US20030004107A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF
; FILE REFERENCE: NEURAL THREAD PROTEINS
; CURRENT APPLICATION NUMBER: US/10/146,130
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1:
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(1139)
US-10-146-130-1

Query Match 100.0%; Score 1442; DB 15; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTTTGTGAGATGAGTTTTCGCTGTTGCCAGGCTGAGTGAATGGCGAA 60
Db 1 TTTTGTGAGATGAGTTTTCGCTGTTGCCAGGCTGAGTGAATGGCGAA 60

Oy 61 TCTCAGCTACACCGAACCTCCGCGCTCCGGGTTCAAGGATTCCTGCTCAGCCTCC 120
Db 61 TCTCAGCTACACCGAACCTCCGCGCTCCGGGTTCAAGGATTCCTGCTCAGCCTCC 120

Oy 121 CAGTAGCTGGGATTAACAGGATGTGCACCCAGCTCCGCTAATTTTGTATTTT 180
Db 121 CAGTAGCTGGGATTAACAGGATGTGCACCCAGCTCCGCTAATTTTGTATTTT 180

Oy 181 TAGAGATGAGTTTCTCATGTTGGTGCAGGCTGGCTGGAAGTCCGACCTCAGATGATC 240
Db 181 TAGAGATGAGTTTCTCATGTTGGTGCAGGCTGGCTGGAAGTCCGACCTCAGATGATC 240

Oy 241 CCTCCGCTCGGCTCCCAAGTGTCTAGATACAGAGCTGGCCACCATGCCGGCTCTGCTC 300

```
Db 241 ||||| CCTCGGCTGGGCTCCCAAGTGTAGATACAGAGTGGGCGACCAAGCCCGGCTCTGGC 300
Qy 301 TGGCTAATTTTGTGTAGAAACAGAGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360
Db 301 ||||| TGGCTAATTTTGTGTAGAAACAGAGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360
Qy 361 TCAGCAGTCCAGCTGCTGAGCTCCCAAGTGTGTGGGATTAAGGCGTGGCAGCCGTGC 420
Db 361 TCAGCAGTCCAGCTGCTGAGCTCCCAAGTGTGTGGGATTAAGGCGTGGCAGCCGTGC 420
Qy 421 CTGGCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
Db 421 CTGGCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
Qy 481 TGCAGTGTGTATCAGAGCTCAGCTCAGCTTCAACTCTGAGATCAGACATCTCTCG 540
Db 481 TGCAGTGTGTATCAGAGCTCAGCTCAGCTTCAACTCTGAGATCAGACATCTCTCG 540
Qy 541 CCTGAGCTCCCAAGTAGGCGGAGCCCAAGATCCACCTACCTACCTGCTATTTT 600
Db 541 CCTGAGCTCCCAAGTAGGCGGAGCCCAAGATCCACCTACCTACCTGCTATTTT 600
Qy 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
Db 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
Qy 661 GGGCAATCTTGGCTACTGCAACCTGCTCCCGGGTTCAGATTTCTCCGCCCCA 720
Db 661 GGGCAATCTTGGCTACTGCAACCTGCTCCCGGGTTCAGATTTCTCCGCCCCA 720
Qy 721 GCTCTGAGTAGCTGAGCTACAGGCGCCACAGCGCTGCTGCTATTTTATTTTATTTT 780
Db 721 GCTCTGAGTAGCTGAGCTACAGGCGCCACAGCGCTGCTGCTATTTTATTTTATTTTATTTT 780
Qy 781 TAGTAGAGTAGGAGGTTTACAGATGTGCGAGGTTGATGTGATGTGATGTGATGTGATGTG 840
Db 781 TAGTAGAGTAGGAGGTTTACAGATGTGCGAGGTTGATGTGATGTGATGTGATGTGATGTG 840
Qy 841 TGCCTGCTGGGCTCCCAAGTGTCTGGGATTAAGGCGTGGAGCCAGCCCGGCTTA 900
Db 841 TGCCTGCTGGGCTCCCAAGTGTCTGGGATTAAGGCGTGGAGCCAGCCCGGCTTA 900
Qy 901 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 960
Db 901 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 960
Qy 961 GGGCAATCTGCTGCTGCAACCTGCTCCCGGGTTCAGATTTCTCCGCCCCA 1020
Db 961 GGGCAATCTGCTGCTGCAACCTGCTCCCGGGTTCAGATTTCTCCGCCCCA 1020
Qy 1021 GCTCTGAGTAGCTGAGCTACAGGCGCCACAGCGCTGCTGCTATTTTATTTTATTTT 1080
Db 1021 GCTCTGAGTAGCTGAGCTACAGGCGCCACAGCGCTGCTGCTATTTTATTTTATTTTATTTT 1080
Qy 1081 TCATTAGAGGCGGGTTCACCAATATTTTCTGAGGTTGCTCAAACTCTGAGCTAGGT 1140
Db 1081 TCATTAGAGGCGGGTTCACCAATATTTTCTGAGGTTGCTCAAACTCTGAGCTAGGT 1140
Qy 1141 GAGCAGCTGCTGAGCTTCCAAAGTGTGGGATTAAGGCGTGGAGCCAGCTCAGCCAG 1200
Db 1141 GAGCAGCTGCTGAGCTTCCAAAGTGTGGGATTAAGGCGTGGAGCCAGCTCAGCCAG 1200
Qy 1201 CGGCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1260
Db 1201 CGGCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1260
Qy 1261 GGTCTCAAACTCTGCTGATGCAATCTTCCAAATAGAGCAGCAAAACAGCCAGCTGTA 1320
Db 1261 GGTCTCAAACTCTGCTGATGCAATCTTCCAAATAGAGCAGCAAAACAGCCAGCTGTA 1320
Qy 1321 CATTTTAAACAGTTACATCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1380
Db 1321 CATTTTAAACAGTTACATCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1380
```

```
Db 1321 CATTTTAAACAGTTACATCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1380
Qy 1381 CAAACCTGCAATTCAGTAGTAACAGAGTTCTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 1440
Db 1381 CAAACCTGCAATTCAGTAGTAACAGAGTTCTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 1440
Qy 1441 CA 1442
Db 1441 CA 1442

RESULT 6
US-10-092-934-1
; Sequence 1, Application US/10092934
; Publication No. US20030054990A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHODS OF USING NEURAL THREAD PROTEINS TO TREAT TUMORS
; TITLE OF INVENTION: AND CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 018792-0199
; CURRENT APPLICATION NUMBER: US/10/092, 934
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/273,957
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO. 1
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)...(1139)
US-10-092-934-1

Query Match 100.0%; Score 1442; DB 15; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 60
Db 1 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 60
Qy 61 TCTGAGTCCAGGAGCAACCTGCTCCCGGGTTCAGAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 61 TCTGAGTCCAGGAGCAACCTGCTCCCGGGTTCAGAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Qy 121 CAGTAGCTGGGATTAAGGAGTGTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 CAGTAGCTGGGATTAAGGAGTGTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 181 TAGAGATGAGTTTCTCATGTGTGTGAGGCTGTGCTGCAACTCCGACCTCAGATGATC 240
Db 181 TAGAGATGAGTTTCTCATGTGTGTGAGGCTGTGCTGCAACTCCGACCTCAGATGATC 240
Qy 241 CCTCGCTGGGCTCCCAAGTGTGATTAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 CCTCGCTGGGCTCCCAAGTGTGATTAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy 301 TGGCTAATTTTGTGTAGAAACAGAGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360
Db 301 TGGCTAATTTTGTGTAGAAACAGAGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360
Qy 361 TCAGCAGTCCAGCTGCTGAGCTCCCAAGTGTGTGGGATTAAGGCGTGGAGCCAGCTGTA 420
Db 361 TCAGCAGTCCAGCTGCTGAGCTCCCAAGTGTGTGGGATTAAGGCGTGGAGCCAGCTGTA 420
Qy 421 CTGGCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
Db 421 CTGGCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
Qy 481 TGCAGTGTGTATCAGAGCTCAGCTCAGCTTCAACTCTGAGATCAGACATCTCTCG 540
Db 481 TGCAGTGTGTATCAGAGCTCAGCTCAGCTTCAACTCTGAGATCAGACATCTCTCG 540
```

```

|||||
481 TCAGAGGTGTATCATCAGCTCAGCTCAGCTTCACTCTGAGATCAAGCATCTCTG 540
541 CCGACGCTCCCAAGAGTGGGACCAAGACATGACACATACATCTGGTATTTTAA 600
541 CCGACGCTCCCAAGAGTGGGACCAAGACATGACACATACATCTGGTATTTTAA 600
601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
661 GCGCAATCTTGGCTCAGTCAACCTCTGCTCCGCGGTCAAGTATTTCTCTGCCCA 720
661 GCGCAATCTTGGCTCAGTCAACCTCTGCTCCGCGGTCAAGTATTTCTCTGCCCA 720
721 GCGCTCTGAGTACAGTGGGACCTACAGGCGCCACACAGCTAGTATTTTGTATTT 780
721 GCGCTCTGAGTACAGTGGGACCTACAGGCGCCACACAGCTAGTATTTTGTATTT 780
781 TAGTAGAGATGGGATTCACCATGTTGCGAGGTTGATCTGTGACCTTGATGATC 840
781 TAGTAGAGATGGGATTCACCATGTTGCGAGGTTGATCTGTGACCTTGATGATC 840
841 TGCCTGCTCGGCTCCCAAGTGTGGATTAACAGGCGTGAGCCACACGCGCGCTTA 900
841 TGCCTGCTCGGCTCCCAAGTGTGGATTAACAGGCGTGAGCCACACGCGCGCTTA 900
901 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 960
901 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 960
961 GCGCAATCTGCGCTCAGTCAACCTCTGCTCCGCGGTCAAGCATTTCTCTGCTCA 1020
961 GCGCAATCTGCGCTCAGTCAACCTCTGCTCCGCGGTCAAGCATTTCTCTGCTCA 1020
1021 GCGCTCTGAGTACAGTGGGACCTACAGGCGCCACACAGCTAGTATTTTGTATTT 1080
1021 GCGCTCTGAGTACAGTGGGACCTACAGGCGCCACACAGCTAGTATTTTGTATTT 1080
1081 TCATTAGAGAGGCGGTTTACCATATTTTGTCAAGGCGGTCTCAAACTCTGACCTAGGT 1140
1081 TCATTAGAGAGGCGGTTTACCATATTTTGTCAAGGCGGTCTCAAACTCTGACCTAGGT 1140
1141 GACCCACCTGCGCTCAGCTTCCAAAGTGTGGATTAACAGGCGTGAGCCACCTCAGCCAG 1200
1141 GACCCACCTGCGCTCAGCTTCCAAAGTGTGGATTAACAGGCGTGAGCCACCTCAGCCAG 1200
1201 CCGGCTAATTTAGATTAATAAATATGTACCAATGGGCGCTCTGTATTTGCGCAGGCT 1260
1201 CCGGCTAATTTAGATTAATAAATATGTACCAATGGGCGCTCTGTATTTGCGCAGGCT 1260
1261 GGTCTCAAACTCTGCTCAGTCAACCTCTTCCAAATGAGCCACACAGCCGCGCTCA 1320
1261 GGTCTCAAACTCTGCTCAGTCAACCTCTTCCAAATGAGCCACACAGCCGCGCTCA 1320
1321 CATTTTATTAACAGTATCATCTTTATTTATTAAGTATTAAGTATTAAGTATTAAGTAT 1380
1321 CATTTTATTAACAGTATCATCTTTATTTATTAAGTATTAAGTATTAAGTATTAAGTAT 1380
1381 CAACCTGCAAAATAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 1440
1381 CAACCTGCAAAATAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 1440
1441 CA 1442
1441 CA 1442

```

RESULT 7
 US-10-153-334-53
 ; Sequence 53, Application US/10153334
 ; Publication No. US20030096350A1
 ; GENERAL INFORMATION:

```

; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; FILE REFERENCE: 59003-000006
; CURRENT APPLICATION NUMBER: US/10/153,334
; PRIOR APPLICATION NUMBER: 2002-05-24
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO: 53
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(1139)
US-10-153-334-53

Query Match      100.0%; Score 1442; DB 15; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 60
Db 1 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 60
QY 61 TCTCAGCTCAGCGCAACCTTCCGCTCCGCGGTCAAGCATTTCTCTGCTCAGCTCCC 120
Db 61 TCTCAGCTCAGCGCAACCTTCCGCTCCGCGGTCAAGCATTTCTCTGCTCAGCTCCC 120
QY 121 CAGTACCTGGGATTAAGCATGTCACACAGCTGCGGTAAATTTTATTTTATTTTATTTG 180
Db 121 CAGTACCTGGGATTAAGCATGTCACACAGCTGCGGTAAATTTTATTTTATTTTATTTG 180
QY 121 CAGTACCTGGGATTAAGCATGTCACACAGCTGCGGTAAATTTTATTTTATTTTATTTG 180
Db 121 CAGTACCTGGGATTAAGCATGTCACACAGCTGCGGTAAATTTTATTTTATTTTATTTG 180
QY 181 TAGAGATGAGTTCATGTTGTGTGACAGCTGCTCGAATCTCCGACCTCAGATGATC 240
Db 181 TAGAGATGAGTTCATGTTGTGTGACAGCTGCTCGAATCTCCGACCTCAGATGATC 240
QY 181 TAGAGATGAGTTCATGTTGTGTGACAGCTGCTCGAATCTCCGACCTCAGATGATC 240
Db 181 TAGAGATGAGTTCATGTTGTGTGACAGCTGCTCGAATCTCCGACCTCAGATGATC 240
QY 241 CTTGCTGCTGCGCTTCCAAAGTGTGATTAAGAGCTGCGCCACATGCGCGCTCTGCC 300
Db 241 CTTGCTGCTGCGCTTCCAAAGTGTGATTAAGAGCTGCGCCACATGCGCGCTCTGCC 300
QY 241 CTTGCTGCTGCGCTTCCAAAGTGTGATTAAGAGCTGCGCCACATGCGCGCTCTGCC 300
Db 241 CTTGCTGCTGCGCTTCCAAAGTGTGATTAAGAGCTGCGCCACATGCGCGCTCTGCC 300
QY 301 TGGCTAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCAAGCTGCTCTGAGC 360
Db 301 TGGCTAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCAAGCTGCTCTGAGC 360
QY 301 TGGCTAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCAAGCTGCTCTGAGC 360
Db 301 TGGCTAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCAAGCTGCTCTGAGC 360
QY 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGATTAACAGGCGTGAGCCGCTGC 420
Db 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGATTAACAGGCGTGAGCCGCTGC 420
QY 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGATTAACAGGCGTGAGCCGCTGC 420
Db 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGATTAACAGGCGTGAGCCGCTGC 420
QY 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
Db 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
QY 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
Db 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
QY 481 TGGAGTGTGATTCACAGCTCAGCTCAGCTTCAACTCTCTGAGATCAAGCATCTCTGCG 540
Db 481 TGGAGTGTGATTCACAGCTCAGCTCAGCTTCAACTCTCTGAGATCAAGCATCTCTGCG 540
QY 481 TGGAGTGTGATTCACAGCTCAGCTCAGCTTCAACTCTCTGAGATCAAGCATCTCTGCG 540
Db 481 TGGAGTGTGATTCACAGCTCAGCTCAGCTTCAACTCTCTGAGATCAAGCATCTCTGCG 540
QY 541 CCGACGCTCCCAAGAGTGGGACCAAGACATGACACATACATCTGGTATTTTAA 600
Db 541 CCGACGCTCCCAAGAGTGGGACCAAGACATGACACATACATCTGGTATTTTAA 600
QY 541 CCGACGCTCCCAAGAGTGGGACCAAGACATGACACATACATCTGGTATTTTAA 600
Db 541 CCGACGCTCCCAAGAGTGGGACCAAGACATGACACATACATCTGGTATTTTAA 600
QY 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
Db 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
QY 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
Db 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
QY 661 GCGCAATCTTGGCTCAGTCAACCTCTGCTCCGCGGTCAAGTATTTCTCTGCCCA 720
Db 661 GCGCAATCTTGGCTCAGTCAACCTCTGCTCCGCGGTCAAGTATTTCTCTGCCCA 720
QY 661 GCGCAATCTTGGCTCAGTCAACCTCTGCTCCGCGGTCAAGTATTTCTCTGCCCA 720
Db 661 GCGCAATCTTGGCTCAGTCAACCTCTGCTCCGCGGTCAAGTATTTCTCTGCCCA 720
QY 721 GCGCTCTGAGTACAGTGGGACCTACAGGCGCCACACAGCTAGTATTTTGTATTTT 780
Db 721 GCGCTCTGAGTACAGTGGGACCTACAGGCGCCACACAGCTAGTATTTTGTATTTT 780

```

DB 721 GCTCTGAGTAGTGGAGCTACAGGGCCCAACGCTGATATTTTGTGATTTT 780
OY 781 TAGTAGAGATGGGTGACATGTGCGCAGGTGATCTGATCTCTGAGACTTGTGATC 840
DB 781 TAGTAGAGATGGGTGACATGTGCGCAGGTGATCTGATCTCTGAGACTTGTGATC 840
OY 841 TGCCTGCTCGGCTCCCAAGTGTGGGATTCACAGGCGTGAACCCGCGGCTTA 900
DB 841 TGCCTGCTCGGCTCCCAAGTGTGGGATTCACAGGCGTGAACCCGCGGCTTA 900
OY 901 TTTTAAATTTTGTGTTGTAATGAAATCTCACTGTGTACCCAGGCTGAGATGCAAT 960
DB 901 TTTTAAATTTTGTGTTGTAATGAAATCTCACTGTGTACCCAGGCTGAGATGCAAT 960
OY 961 GGCCTAATCTCGGCTCACTGCAACCTGTGCTCCGCGGCTCAAGGATTCCTGCTCA 1020
DB 961 GGCCTAATCTCGGCTCACTGCAACCTGTGCTCCGCGGCTCAAGGATTCCTGCTCA 1020
OY 1021 GCTCCCAAGCAGCTGGGATTCAGGCGACCTGCGCACACCCGCTAATTTTGTATTT 1080
DB 1021 GCTCCCAAGCAGCTGGGATTCAGGCGACCTGCGCACACCCGCTAATTTTGTATTT 1080
OY 1081 TCATTAGAGGCGGGGTTTCACCAATTTTGTGAGGCTGTCTCAACTCTGACCTCAGAT 1140
DB 1081 TCATTAGAGGCGGGGTTTCACCAATTTTGTGAGGCTGTCTCAACTCTGACCTCAGAT 1140
OY 1141 GACCCAGCTGCTCAGCTTCCAAAGTGTGGGATTCAGGCGTGAACCTGACCTCAGAT 1200
DB 1141 GACCCAGCTGCTCAGCTTCCAAAGTGTGGGATTCAGGCGTGAACCTGACCTCAGAT 1200
OY 1201 CGGGCTAATTTAGATTAATAAATAATGTAGCAATGGGGGCTGTGCTAGTGGCCAGGCT 1260
DB 1201 CGGGCTAATTTAGATTAATAAATAATGTAGCAATGGGGGCTGTGCTAGTGGCCAGGCT 1260
OY 1261 GGTTCACAACTTCTGCTCATGATCTCTCAATGAGCAGACACCCAGGCTGATCA 1320
DB 1261 GGTTCACAACTTCTGCTCATGATCTCTCAATGAGCAGACACCCAGGCTGATCA 1320
OY 1321 CATTTTAAACAGTTACATCTTATTTAGTACTAGTAAGTAATCAATAAACAATGT 1380
DB 1321 CATTTTAAACAGTTACATCTTATTTAGTACTAGTAAGTAATCAATAAACAATGT 1380
OY 1381 CAACCTGCAATTCAGTAGTAACAGAGTCTTTTAACTTTTAAACAAGCTTTAGAG 1440
DB 1381 CAACCTGCAATTCAGTAGTAACAGAGTCTTTTAACTTTTAAACAAGCTTTAGAG 1440
OY 1441 CA 1442
DB 1441 CA 1442

RESULT 8
US-10-198-069-48
: Sequence 48, Application US/10198069
: Publication No. US20030096756A1
: GENERAL INFORMATION:
: APPLICANT: AVERBACK, PAUL
: TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
: TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
: FILE REFERENCE: 59003.000009
: CURRENT APPLICATION NUMBER: US/10/198.069
: PRIOR FILING DATE: 2002-07-19
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/306,161
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/331,477
: PRIOR FILING DATE: 2001-11-16
: NUMBER OF SEQ ID NOS: 48
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 48

LENGTH: 1442
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (15)..(1139)
US-10-198-069-48
Query Match 100.0%; Score 1442; DB 15; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTTTAAATTTTGTGTTGTAATGAAATCTCACTGTGTACCCAGGCTGAGATGCAAT 60
DB 1 TTTTAAATTTTGTGTTGTAATGAAATCTCACTGTGTACCCAGGCTGAGATGCAAT 60
OY 61 TCTCAGCTCAGGCACTCCGCTCCGCGGTTCAAGGATTCCTGCTCAGCTCC 120
DB 61 TCTCAGCTCAGGCACTCCGCTCCGCGGTTCAAGGATTCCTGCTCAGCTCC 120
OY 121 CAGTAGCTGGGATTCACAGCATGTGACCCAGCTGGCTAATTTTGTATTTTGTAG 180
DB 121 CAGTAGCTGGGATTCACAGCATGTGACCCAGCTGGCTAATTTTGTATTTTGTAG 180
OY 181 TAGAGATGAGATTTCTCATGATGTGCTAGGCTGTCTGCAATCCGACCTCAGATGATC 240
DB 181 TAGAGATGAGATTTCTCATGATGTGCTAGGCTGTCTGCAATCCGACCTCAGATGATC 240
OY 241 CCGCGCTCGGCTCCCAAGTGTGATACAGAGCTGGGACAGTCCGCGGCTGCGC 300
DB 241 CCGCGCTCGGCTCCCAAGTGTGATACAGAGCTGGGACAGTCCGCGGCTGCGC 300
OY 301 TGGCTAATTTTGTGTTGTAATGAAATCTCACTGTGTGCGCAAGCTGTCTCTGAGC 360
DB 301 TGGCTAATTTTGTGTTGTAATGAAATCTCACTGTGTGCGCAAGCTGTCTCTGAGC 360
OY 361 TCAGAGCTCAGCTGCTCAGCTCCCAAGTGTGAGGATTCAGGCGTGAACCTGCTG 420
DB 361 TCAGAGCTCAGCTGCTCAGCTCCCAAGTGTGAGGATTCAGGCGTGAACCTGCTG 420
OY 421 CTGGGCTTTTATTTTATTTTATTTTATTTTAAAGACAGTGTCTCCACTTACCAGGATGAG 480
DB 421 CTGGGCTTTTATTTTATTTTATTTTATTTTAAAGACAGTGTCTCCACTTACCAGGATGAG 480
OY 481 TGGAGTGGTGGATTCACAGCTGCACTGAGCTTCAACTCTCTGAGATCAAGATCTCTG 540
DB 481 TGGAGTGGTGGATTCACAGCTGCACTGAGCTTCAACTCTCTGAGATCAAGATCTCTG 540
OY 541 CCTCAGCTCCCAAGTGTGAGCAGCAAGACATGCACTACCTACCTGCTGCTAATTTT 600
DB 541 CCTCAGCTCCCAAGTGTGAGCAGCAAGACATGCACTACCTACCTGCTGCTAATTTT 600
OY 601 TTTTAAATTTTGTGTTGTAATGAAATCTCACTGTGTACCCAGGCTGAGATGCAAT 660
DB 601 TTTTAAATTTTGTGTTGTAATGAAATCTCACTGTGTACCCAGGCTGAGATGCAAT 660
OY 661 GGGCAATCTGGGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 720
DB 661 GGGCAATCTGGGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 720
OY 721 GCTCTGAGTAGTGGGACTACAGGCGCCACACGCTGACTAATTTTGTATTTT 780
DB 721 GCTCTGAGTAGTGGGACTACAGGCGCCACACGCTGACTAATTTTGTATTTT 780
OY 781 TAGTAGAGATGGGTGACATGTGCGCAGGTGATCTGATCTCTGAGACTTGTGATC 840
DB 781 TAGTAGAGATGGGTGACATGTGCGCAGGTGATCTGATCTCTGAGACTTGTGATC 840
OY 841 TGCCTGCTCGGCTCCCAAGTGTGGGATTCACAGGCGTGAACCCGCGGCTTA 900
DB 841 TGCCTGCTCGGCTCCCAAGTGTGGGATTCACAGGCGTGAACCCGCGGCTTA 900
OY 901 TTTTAAATTTTGTGTTGTAATGAAATCTCACTGTGTACCCAGGCTGAGATGCAAT 960

Dp	901	TTTTTAATTTTTTGTGTTGTTGAATGGAATCTCACTCTGTTACCGAGGCTGGAGTGCAT	960
QY	961	GGCCAAATCTCGGCTCACTGCAACCTGCGCTCCGGGCTCAAGGATTCCTGTCTCA	1020
Dp	961	GGCCAAATCTCGGCTCACTGCAACCTGCGCTCCGGGCTCAAGGATTCCTGTCTCA	1020
QY	1021	GCCTCCCAAGAGCTGGGATTACGGGGACGTGCCACACACCCGCTAATTTTGTATT	1080
Dp	1021	GCCTCCCAAGAGCTGGGATTACGGGGACGTGCCACACACCCGCTAATTTTGTATT	1080
QY	1081	TCATTAGAGGGGGGTTTCACCATATTTGTACGGCTGGTCTCAAACTCTGACCTCAGT	1140
Dp	1081	TCATTAGAGGGGGGTTTCACCATATTTGTACGGCTGGTCTCAAACTCTGACCTCAGT	1140
QY	1141	GACCACCTGGCTCAGCCTTCCAAAGCTGTGGGATTACAGGCTGAGGACCTACCCAG	1200
Dp	1141	GACCACCTGGCTCAGCCTTCCAAAGCTGTGGGATTACAGGCTGAGGACCTACCCAG	1200
QY	1201	CCGGCTAATTTAGATAAAAAATATGTAGCAATGGGGGCTGTGCTATGTTGCCAGGCT	1260
Dp	1201	CCGGCTAATTTAGATAAAAAATATGTAGCAATGGGGGCTGTGCTATGTTGCCAGGCT	1260
QY	1261	GCTCTCAACCTTCTGGCTTCATGCAATCTCTTCCAAATGACGCACACACCCAGCCAGTCA	1320
Dp	1261	GCTCTCAACCTTCTGGCTTCATGCAATCTCTTCCAAATGACGCACACACCCAGCCAGTCA	1320
QY	1321	CATTTTTTAAACAGTTACATCTTATTTTGTAGTATCTGAAAGTATATCAATTAACATGT	1380
Dp	1321	CATTTTTTAAACAGTTACATCTTATTTTGTAGTATCTGAAAGTATATCAATTAACATGT	1380
QY	1381	CAAACTGCAAAATTCAGTAGTAACAGAGTCTTTTATTAACCTTTTAAACAAGCTTTAGAG	1440
Dp	1381	CAAACTGCAAAATTCAGTAGTAACAGAGTCTTTTATTAACCTTTTAAACAAGCTTTAGAG	1440
QY	1441	CA 1442	
Dp	1441	CA 1442	

RESULT 9
US-10-198-070-125
; Sequence 125, Application, US/10198070
; Publication No. US20030109437A1

```

1  APPLICANT: GEMMELL, JACK
2  TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
3  TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
4  TITLE OF INVENTION: CELLS
5  FILE REFERENCE: 59003.000008
6  CURRENT APPLICATION NUMBER: US/10/198,070
7  CURRENT FILING DATE: 2002-07-19
8  PRIOR APPLICATION NUMBER: 60/306,161
9  PRIOR FILING DATE: 2001-07-19
10 PRIOR APPLICATION NUMBER: 60/306,150
11 PRIOR FILING DATE: 2001-07-19
12 PRIOR APPLICATION NUMBER: 60/731,477
13 PRIOR FILING DATE: 2001-11-16
14 NUMBER OF SEQ ID NOS: 125
15 SOFTWARE: PatentIn Ver. 2.1
16 SEQ ID NO 125
17 LENGTH: 1442
18 TYPE: DNA
19 ORGANISM: Homo sapiens
20 FEATURE:
21 NAME/KEY: CDS
22 LOCATION: (15)..(1139)
23 US-10-198-070-125

```

Query Match	100.0%;	Score 1442;	DB 15;	Length 1442;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1442;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	TTTTTTTTTTTACAGATGGATTTTGCCTCTGTGTGCCACAGCGGAGATGAAAGGGCAA	60
Db	1 TTTTTTTTTTTAGATGAGATTTTGCCTCTGTGTGCCACAGCGGAGATGAAAGGGCAA	60
QY	TTCTACGCTACCGCAACCTCCGCGCTCCCGGGTTTCAAGCGATTTCTCCGTCCAGCTCC	120
Db	61 TCTACGCTACCGCAACCTCCGCGCTCCCGGGTTTCAAGCGATTTCTCCGTCCAGCTCC	120
QY	TTCTACGCTACCGCAACCTCCGCGCTCCCGGGTTTCAAGCGATTTCTCCGTCCAGCTCC	120
Db	61 TCTACGCTACCGCAACCTCCGCGCTCCCGGGTTTCAAGCGATTTCTCCGTCCAGCTCC	120
QY	CAGTAGCTGGGATTTACAGGCATGTGACCCAGCGTGGGCTAATTTTGTATTTT	180
Db	121 CAGTAGCTGGGATTTACAGGCATGTGACCCAGCGTGGGCTAATTTTGTATTTT	180
QY	CAGTAGCTGGGATTTACAGGCATGTGACCCAGCGTGGGCTAATTTTGTATTTT	180
Db	121 CAGTAGCTGGGATTTACAGGCATGTGACCCAGCGTGGGCTAATTTTGTATTTT	180
QY	TAGAGATGGAGATTTCTCATGTGTGGTGCAGCGTGGTCTCGAATCTCCGACCTCAGATGAT	240
Db	181 TAGAGATGGAGATTTCTCATGTGTGGTGCAGCGTGGTCTCGAATCTCCGACCTCAGATGAT	240
QY	TAGAGATGGAGATTTCTCATGTGTGGTGCAGCGTGGTCTCGAATCTCCGACCTCAGATGAT	240
Db	181 TAGAGATGGAGATTTCTCATGTGTGGTGCAGCGTGGTCTCGAATCTCCGACCTCAGATGAT	240
QY	CCTCCGCTCTCGGCTCCCAAAGTGCATGATACAGACTGGCCACCAATGCCCGGCTCGC	300
Db	241 CCTCCGCTCTCGGCTCCCAAAGTGCATGATACAGACTGGCCACCAATGCCCGGCTCGC	300
QY	CCTCCGCTCTCGGCTCCCAAAGTGCATGATACAGACTGGCCACCAATGCCCGGCTCGC	300
Db	241 CCTCCGCTCTCGGCTCCCAAAGTGCATGATACAGACTGGCCACCAATGCCCGGCTCGC	300
QY	TGGCTAATTTTGTGGTAGAAGAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGC	360
Db	301 TGGCTAATTTTGTGGTAGAAGAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGC	360
QY	TGGCTAATTTTGTGGTAGAAGAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGC	360
Db	301 TGGCTAATTTTGTGGTAGAAGAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGC	360
QY	TCAGCAGTCCACCTGCTCAGCGCTCCCAAAGTGTGGGATTCAGGCGTGCAGCGCTGC	420
Db	361 TCAGCAGTCCACCTGCTCAGCGCTCCCAAAGTGTGGGATTCAGGCGTGCAGCGCTGC	420
QY	TCAGCAGTCCACCTGCTCAGCGCTCCCAAAGTGTGGGATTCAGGCGTGCAGCGCTGC	420
Db	361 TCAGCAGTCCACCTGCTCAGCGCTCCCAAAGTGTGGGATTCAGGCGTGCAGCGCTGC	420
QY	CTGGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	480
Db	421 CTGGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	480
QY	CTGGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	480
Db	421 CTGGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	480
QY	TGCAGTGTGTATCATCAGACTACTGACACTTCAACTCTCTGAGATCAAGCATCTCTG	540
Db	481 TGCAGTGTGTATCATCAGACTACTGACACTTCAACTCTCTGAGATCAAGCATCTCTG	540
QY	TGCAGTGTGTATCATCAGACTACTGACACTTCAACTCTCTGAGATCAAGCATCTCTG	540
Db	481 TGCAGTGTGTATCATCAGACTACTGACACTTCAACTCTCTGAGATCAAGCATCTCTG	540
QY	CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGACACATACACTTGCTAATTTT	600
Db	541 CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGACACATACACTTGCTAATTTT	600
QY	CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGACACATACACTTGCTAATTTT	600
Db	541 CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGACACATACACTTGCTAATTTT	600
QY	TTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	660
Db	601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	660
QY	TTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	660
Db	601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	660
QY	GGCGCAATCTTGGCTACTGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCCCA	720
Db	661 GGCGCAATCTTGGCTACTGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCCCA	720
QY	GGCGCAATCTTGGCTACTGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCCCA	720
Db	661 GGCGCAATCTTGGCTACTGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCCCA	720
QY	GCTCTCTGATGTAGCTGGGACCTACAGCGGCCCAACAGCGCTAGCTAATTTTGTATTTT	780
Db	721 GCTCTCTGATGTAGCTGGGACCTACAGCGGCCCAACAGCGCTAGCTAATTTTGTATTTT	780
QY	GCTCTCTGATGTAGCTGGGACCTACAGCGGCCCAACAGCGCTAGCTAATTTTGTATTTT	780
Db	721 GCTCTCTGATGTAGCTGGGACCTACAGCGGCCCAACAGCGCTAGCTAATTTTGTATTTT	780
QY	TAGTAGAGATGGGGTTTACACATGTTTCGCCAGGTGTGATCTGTGACCTTGTGATC	840
Db	781 TAGTAGAGATGGGGTTTACACATGTTTCGCCAGGTGTGATCTGTGACCTTGTGATC	840
QY	TAGTAGAGATGGGGTTTACACATGTTTCGCCAGGTGTGATCTGTGACCTTGTGATC	840
Db	781 TAGTAGAGATGGGGTTTACACATGTTTCGCCAGGTGTGATCTGTGACCTTGTGATC	840
QY	TGCGTCTCGGCTCGGCTCCCAAAGTGTGGGATTCAGAGCGTGAAGCCACACAGCCGGCTTA	900
Db	841 TGCGTCTCGGCTCGGCTCCCAAAGTGTGGGATTCAGAGCGTGAAGCCACACAGCCGGCTTA	900
QY	TGCGTCTCGGCTCGGCTCCCAAAGTGTGGGATTCAGAGCGTGAAGCCACACAGCCGGCTTA	900
Db	841 TGCGTCTCGGCTCGGCTCCCAAAGTGTGGGATTCAGAGCGTGAAGCCACACAGCCGGCTTA	900
QY	TTTTTAAATTTTGTGTTTGTGAATTTGGGATCTCTGTACCCAGGCTGGAGTGCAT	960
Db	901 TTTTAAATTTTGTGTTTGTGAATTTGGGATCTCTGTACCCAGGCTGGAGTGCAT	960
QY	TTTTTAAATTTTGTGTTTGTGAATTTGGGATCTCTGTACCCAGGCTGGAGTGCAT	960
Db	901 TTTTAAATTTTGTGTTTGTGAATTTGGGATCTCTGTACCCAGGCTGGAGTGCAT	960
QY	GGCGCAATCTCGGCTACTGCAACCTCTGCTCCGGGCTCAAGCGATTTCTCTGCTCA	1020
Db	961 GGCGCAATCTCGGCTACTGCAACCTCTGCTCCGGGCTCAAGCGATTTCTCTGCTCA	1020
QY	GGCGCAATCTCGGCTACTGCAACCTCTGCTCCGGGCTCAAGCGATTTCTCTGCTCA	1020
Db	961 GGCGCAATCTCGGCTACTGCAACCTCTGCTCCGGGCTCAAGCGATTTCTCTGCTCA	1020
QY	GCTCTCCCAAGCAGCTGGGATTCAGGGCAGCTGCACACACCCCGCTAATTTTGTATTT	1080
Db	1021 GCTCTCCCAAGCAGCTGGGATTCAGGGCAGCTGCACACACCCCGCTAATTTTGTATTT	1080
QY	GCTCTCCCAAGCAGCTGGGATTCAGGGCAGCTGCACACACCCCGCTAATTTTGTATTT	1080
Db	1021 GCTCTCCCAAGCAGCTGGGATTCAGGGCAGCTGCACACACCCCGCTAATTTTGTATTT	1080

QY 1081 TCATTAGAGCGGGGTTTACCATTTGTACAGCTGCTGCTTCAAACTCTGACCTCAGGT 1140
 Db 1081 TCATTAGAGCGGGGTTTACCATTTGTGTAGGCTGCTGCTCAAACTCTGACCTCAGGT 1140
 QY 1141 GACCCACCTGCTGCTGCTTCCAAAGTCTGGAATTCAGGCGTACAGCCAGCCAG 1200
 Db 1141 GACCCACCTGCTGCTGCTTCCAAAGTCTGGAATTCAGGCGTACAGCCAGCCAG 1200
 QY 1201 CCGGCTAATTTAGATTAATAAATATGTAGCAATGGGGGCTTGTGATGTTGCCAGGCT 1260
 Db 1201 CCGGCTAATTTAGATTAATAAATATGTAGCAATGGGGGCTTGTGATGTTGCCAGGCT 1260
 QY 1261 GGTCTCAACTTCTGCTTCAATGCAATCTTCCAAATGAGCCAGCAACACCCAGCCAGTCA 1320
 Db 1261 GGTCTCAACTTCTGCTTCAATGCAATCTTCCAAATGAGCCAGCAACACCCAGCCAGTCA 1320
 QY 1321 CATTTTAAACAGTTTCACTTATTTTATTTAGTATCTAGAAATTAATTAACATGAT 1380
 Db 1321 CATTTTAAACAGTTTCACTTATTTTATTTAGTATCTAGAAATTAATTAACATGAT 1380
 QY 1381 CAACCTGCAATTCAGTACTTAACAGAGTCTTTTAACTTTTAAACAAAGCTTTAGAG 1440
 Db 1381 CAACCTGCAATTCAGTACTTAACAGAGTCTTTTAACTTTTAAACAAAGCTTTAGAG 1440
 QY 1441 CA 1442
 Db 1441 CA 1442

RESULT 10

US-09-964-666-4
 Sequence 4, Application US/09964666
 Patent No. US20020104108A1
 GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
 Wands, Jack R.

TITLE OF INVENTION: Transgenic Animals and Cell Lines for
 Screening Drugs Effective for the Treatment or Prevention
 of Alzheimer's Disease

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
 STREET: 1100 New York Ave., Suite 600
 CITY: Washington
 STATE: DC

COUNTRY: USA
 ZIP: 20005-3934

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/964,666
 FILING DATE: 28-Sep-2001
 CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert M.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0609,4370000

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2540

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1418 base pairs
 TYPE: nucleic acid

STRANDEDNESS: both
 TOPOLOGY: both

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-964-666-4

Query Match 84.8%; Score 1223.4; DB 11; Length 1418;
 Best Local Similarity 96.9%; Pred. No. 0;
 Matches 1375; Conservative 0; Mismatches 31; Indels 13; Gaps 12;

QY 2 TTTTATTTTGTAGATGAGTTTGTGCTCTTTTCCAGGCTGAGTGAATGGCGCAAT 61
 Db 1 TTTTATTTTGTAGATGAGTTTGTGCTCTTTTCCAGGCTGAGTGAATGGCGCAAT 60
 QY 62 CTCACCTACCGCAACCTCCGCTCCGGGTTCAAGGATTTCTCCCTCACCCTCCC 121
 Db 61 CTCACCTACCGCAACCTCCGCTCCGGGTTCAAGGATTTCTCCCTCACCCTCCC 120
 QY 122 AGTA-GCTGGATTAAGAGGATGTGCAACCACTGCTGGTAAATTTTGTATTTTGTAG 180
 Db 121 AGTAGGCTGGATTAAGAGGATGTGCA-CGACCTGGCTAAATTTTGTATTTTGTAG 179
 QY 181 TAGAGATGAGTTTCTCCATGTTGTCAGCTGCTGCTCAACT-CCGACCTGATGATG 240
 Db 180 TAGAGATGAGTTTCTCCATGTTGTCAGCTGCTGCTCAACT-CCGACCTGATGATG 238
 QY 241 CCTCGCTCTGGGCTCCCAAGTGTATGATACAGACTGCGCACCATGCCCCG-CTCTGC 299
 Db 239 CTCCGCTCTGGGCTCCCAAGTGTATGATACAGACTGCGCACCATGCCCCGCTCTGC 298
 QY 300 CTGCTAATTTTGTGTAGAAACAGGTTTCACTGATGCTGCAAGCTGCTCTCTAG 359
 Db 299 CTGCTAATTTTGTGTAGAAACAGGTTTCACTGATGCTGCAAGCTGCTCTCTAG 358
 QY 360 CTCACGAGTCACACCTGCTCCAGCTCCCAAGTGTGGATTTACAGGCTGACGCGCTG 419
 Db 359 CTCACGAGTCACACCTGCTCCAGCTCCCAAGTGTGGATTTACAGGCTGACGCGCTG 418
 QY 420 CTTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 479
 Db 419 CTTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 478
 QY 480 GTGACGTGTGTGATCAGACGCTCAGCTGACGCTTCACTGATGATGATGATGATG 539
 Db 479 GTGACGTGTGTGATCAGACGCTCAGCTGACGCTTCACTGATGATGATGATGATG 537
 QY 540 GCCTCAGCTCC-AGTAGCTGAGCAAGACATGACATGACATGACATGATTTT 598
 Db 538 GCCTCAGCTCCCAAGTAGCTGAGCAAGACATGACATGACATGACATGATTTT 597
 QY 599 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 658
 Db 598 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 657
 QY 659 GTGGGCAATCTTGGCTCAGTCAACCTGCTGCTCCGGGTTCAAGTATTTCTGCTGCC 718
 Db 658 GTGGGCAATCTTGGCTCAGTCAACCTGCTGCTCCGGGTTCAAGTATTTCTGCTGCC 717
 QY 719 CAGCTCTGAGTGTGAGCTGAGCTGAGGCGCCACAGCGCTGATTTTATTTTATTTT 778
 Db 718 CAGCTCTGAGTGTGAGCTGAGCTGAGGCGCCACAGCGCTGATTTTATTTTATTTT 777
 QY 779 TTTAGTAGAGATGGG-TTTCCATGTTGCTGAGTTGAT-CTTGTCTCTGACCTGT 836
 Db 778 TTTAGTAGAGATGGGTTTCCATGTTGCTGAGTTGAT-CTTGTCTCTGACCTGT 837
 QY 837 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
 Db 838 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
 QY 896 GCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 955
 Db 898 GCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 957
 QY 956 GCAATGGCCAAATCTGCTCAGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015
 Db 958 GCAATGGCCAAATCTGCTCAGTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1017
 QY 1016 TCTCAGCTTCCCAAGAGCTGGGATTTACGGGCACTGCTCCACACACCCCGCTAATTTT 1075

Db 1018 TCTCAGCTCCCAAGCAGCTGGGATTTACGGGACCTG-CACCAACCCCGCTAATTTTGG 1076
Qy 1076 TATTTTCATATAGAGGGGGGTTTCACATATTTGTGTCAGGCTGGTCTCAAACTCTGACCT 1135
Db 1077 TATTTTCATATAGAGGGGGGTTTCACATATTTGTGTCAGGCTGGTCTCAAACTCTGACCT 1136
Qy 1136 CAGGTGACCCAGCTGCTCAGGCTTCGCAAAAGTGGGATTTACAGGAGGAGCAACCTCA 1195
Db 1137 CAGGTGACCCAGCTGCTCAGGCTTCGCAAAAGTGGGATTTACAGGAGGAGCAACCTCA 1194
Qy 1196 CCGGAGCGGCTATTTAGATATAAATAATGTACATGGGGGCTCTGTATGTTGCC 1255
Db 1195 CCGGAGCGGCTATTTAGATATAAATAATGTACATGGGGGCTCTGTATGTTGCC 1254
Qy 1256 AGGCTGGTCTCAAACTGCTGGCTTCATGCAATCTTCCAAATGAGCCAGCAACCCAGCC 1315
Db 1255 AGGCTGGTCTCAAACTGCTGGCTTCATGCAATCTTCCAAATGAGCCAGCAACCCAGCC 1314
Qy 1316 AGTCACATTTTAAACAGTATCATCTTATTTAGTATACATGAAAGTATACATATA 1375
Db 1315 AGTCACA-TTTTAAACAGTATCATCTTATTTAGTATACATGAAAGTATACATATA 1373
Qy 1376 CATGTCAAACTGCAATTCAGTATGATACAGAGTCTTT 1414
Db 1374 ATGGCGAGCTGCAAAATTCAGTATGATACAGAGTCTTTT 1412

RESULT 12
US-09-964-667-4
; Sequence 4, Application US/09964667
; Publication No. US2003033621A1
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; Mand, Jack R.
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for
; Screening Drugs Effective for the Treatment or Prevention
; of Alzheimer's Disease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/964,667
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0609.4370000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-964-667-4

Query Match 84.8%; Score 1223.4; Db 12; Length 1418;
Best Local Similarity 96.9%; Pred. No. 0;

Matches 1375; Conservative 0; Mismatches 31; Indels 13; Gaps 12;
Qy 2 TTTTATTTTATGATGAGATTTTCGCTCTGTTGCCAGGCTGAGTGCATATGGCCANT 61
Db 1 TTTTATTTTATGATGAGATTTTCGCTCTGTTGCCAGGCTGAGTGCATATGGCCANT 60
Qy 62 CTCAGGCTCACCGCAACCTCGGCTCCCGGGTTCAAGGATTTCTGCTCAGCTCCGCC 121
Db 61 CTCAGGCTCACCGCAACCTCGGCTCCCGGGTTCAAGGATTTCTGCTCAGCTCCGCC 120
Qy 122 AGTA-GCTGGGATTTACAGCATGTGACCCAGCTCGGCTAATTTTATTTTATG 180
Db 121 AGTAAAGGAGATTTACAGCATGTGACCCAGCTCGGCTAATTTTATTTTATG 179
Qy 181 TAGAGATGAGATTTCTCATGTTGTGTCAGGCTGTCTGCAACTCCGACCTCAGATATC 240
Db 180 TAGAGATGAGATTTCTCATGTTGTGTCAGGCTGTCTGCAACT-CCGACCTCAGATATC 238
Qy 241 CCGCGCTCGGCGCTCCCAAGTGTAGATACAGGATGGCCAGATGCCCGG-CTCTGC 299
Db 239 CTCCGCTCGGCGCTCCCAAGTGTAGATACAGGATGGCCAGATGCCCGGCTCTGC 298
Qy 300 CTGGCTAATTTTGTGTAGAAACAGGCTTCACTGATGTGCCAAGCTGTCTCTGAG 359
Db 299 CTGGCTAATTTTGTGTAGAAACAGGCTTCACTGATGTGCCAAGCTGTCTCTGAG 358
Qy 360 CTCAGCAGTCCACCTGCTCAGCTCCCAAGTGTCTGGATTTACAGGCTGTGACGCCGTG 419
Db 359 CTCAGCAGTCCACCTGCTCAGCTCCCAAGTGTCTGGATTTACAGGCTGTGACGCCGTG 418
Qy 420 CCGGAGCTTTTATTTTATTTTATTTTAAAGCAGGCTTCCCATCTTACCCAGATGAA 479
Db 419 CCGGAGCTTTTATTTTATTTTATTTTAAAGCAGGCTTCCCATCTTACCCAGATGAA 478
Qy 480 GTGCAAGTGTGTATGATACAGCTCAGCTGACCTTAATCTGTGATGATACAGATCTCT 539
Db 479 GTGCAAGTGTGTATGATACAGCTCAGCTGACCTTAATCTGTGATGATACAGATCTCT 537
Qy 540 GCCTCAGCTCC-AGATGATGGGACCAAGACATGACACTACACTGAGCTAATTTT 598
Db 538 GCCTCAGCTCC-AGATGATGGGACCAAGACATGACACTACACTGAGCTAATTTT 597
Qy 599 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 658
Db 598 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 657
Qy 659 GTGGGCAATGTGGCTCAGTCACTGCAACTGTGCTCCCGGGTTCAAGTATTTCTGCCC 718
Db 658 GTGGGCAATGTGGCTCAGTCACTGCAACTGTGCTCCCGGGTTCAAGTATTTCTGCCC 717
Qy 719 CAGCTCCTGAGTACGTGAGTACAGGCGCCACACAGCTTACCTAATTTTGTATTT 778
Db 718 CAGCTCCTGAGTACGTGAGTACAGGCGCCACACAGCTTACCTAATTTTGTATTT 777
Qy 779 TTTAGTATGATGAGG-TTCACCATGTTGCCAGGTTGAT-CTTGATGCTGACCTTGT 836
Db 778 TTTAGTATGATGAGGTTTACCATGTTGCCAGGTTGATGATGATGCTTGTACCTTGT 837
Qy 837 CATCTGCTGCTGCTGCTCCCAAGTGTGAGATTTAAGG-CGTGAGCAGCACCCCGG 895
Db 836 CATCTGCTGCTGCTGCTCCCAAGTGTGAGATTTAAGG-CGTGAGCAGCACCCCGG 897
Qy 896 GCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 955
Db 895 GCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 957
Qy 956 GCAATGAGCAATGTGGCTCAGTCAACTGTGCTCCCGGGTCAAGGATTTCTGCG 1015
Db 955 GCAATGAGCAATGTGGCTCAGTCAACTGTGCTCCCGGGTCAAGGATTTCTGCG 1017
Qy 1016 TCTCAGCTCCCAAGCAGCTGGGATTTACGGGACCTGCAACACCCCGCTAATTTTGG 1075
Db 1018 TCTCAGCTCCCAAGCAGCTGGGATTTACGGGACCTG-CACCAACCCCGCTAATTTTGG 1076

QY 1072 TTTGATTTTATTAGAGGGGGTTTACACATTTTGTGAGGCT-GGTCTCAACTGCT 1130
1131 GACCTCAGGTACCCACCTGCTCCTCAGCTTCCCAAGTGTGGGATTTACAGCGGTAGCCA 1190
1129 GACCTCAGGTACCCACCTGCTCCTCAGCTTCCCAAGTGTGGGATTTACAGCGGTAGCCA 1188
QY 1191 CCTCACCAGCGGCTTAATTAGATTAATAAATATGATAGCAATGGGGGCTTGTATGT 1250
1189 CCTCACCAGCGGCTTAATTAGATTAATAAATATGATAGCAATGGGGG--TGTGTATGT 1246
QY 1251 TGGCCAGGCTGTCTCAACTTGTGCTTCATGCAATCTCTTCAATAGCCACCAACCC 1310
1247 TGGCCAGGCTGTCTCAACTTGTGCTTCATGCAATCTCTTCAATAGCCACCAACCC 1306
QY 1311 CAGCCAGTCACATTTTAAACAGTTACATCTTATTTAGTATAGTAAAGTAAATACA 1370
1307 CAGCCAGTCACATTTTAAACAGTTACATCTTATTTAGTATAGTAAAGTAAATACA 1366
QY 1371 ATAAACATGTCAAC 1385
1367 ATAAACATGTCAAC 1381
Db

RESULT 14
US-09-964-412-3
Sequence 3, Application US/09964412
Patent No. US2002129391A1
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
Mands, Jack R.
TITLE OF INVENTION: Transgenic Animals and Cell Lines for
Screening Drugs Effective for the Treatment or Prevention
of Alzheimer's Disease
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
STREET: 1100 New York Ave., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/964,412
FILING DATE: 28-Sep-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Emond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0609,4370000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1381 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: CDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-964-412-3

Query Match 74.9%; Score 1080.2; DB 11; Length 1381;
Best Local Similarity 94.4%; Pred. No. 2,6e-304;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

QY 2 TTTTATTTTATGATGAGTTTCGCTCTTGTGGCCAGGCTGAGTGCATAGGCCAAT 61
1 TTTTATTTTATGATGAGTTTCGCTCTTGTGGCCAGGCTGAGTGCATAGGCCAAT 60
Db

QY 62 CTCACCTACCGCAACCTCGGCTCCGGGTTCAAGGATTCCTGTCCTCAGCTCCGC 121
61 CTCACCTACCGCAACCTCGGCTCCGGGTTCAAGGATTCCTCCTCAGCTCCGC 120
Db

QY 122 AGTACCTGGATTTACAGGATTTGACACCAGCTGGCTAATTTTGTATTTTATTTT 181
121 AGTACCTGGATTTACAGGATTTGACACCAGCTGGCTAATTTTGTATTTTATTTT 179
Db

QY 182 AGAGATGAGTTT--CTCCATGTTGGTACAGGCTGTCTGCACTCCGACCTCAGATGT 239
180 AGAGATGAGTTTAACTCCATGTTGGTACAGGCTGTCTGCACTCCGACCTCAGATGT 239
QY 240 CCTCCGCTCGGCTCCCAAGTGTCT--AGATACAGGATGAGCAGATGCGCGG--CT 295
240 CCTCCGCTCGGCTCCCAAGTGTCTGAGATTTACAGGATGAGCAGATGCGCGGCT 299
Db

QY 296 CTGCTGGCTAATTTTGTGTGTAAGAACAGGTTTCACTGATG--TGGCCAGCTGCTC 354
300 CTGCTGGCTAATTTTGTGTGTAAGAACAGGTTTCACTGATGTTGCCAAGCTGCTC 359
Db

QY 355 CTGAGCTCAAGCAGTCCACCTGCTCAGGCTCCCAAGTGTGGATTTACAGGCTGAG 414
360 CTGAGCTCAAGCAGTCCACCTGCTCAGGCTCCCAAGTGTGGATTTACAGGCTGAG 418
Db

QY 415 CCGTCCCTGGCTTTTATTTATTTTATTTTAAAGACAGGTTGCTCCACTTACCCAG 474
419 CCGTCCCTGGCTTTTATTTATTTTATTTTAAAGACAGGTTGCTCCACTTACCCAG 478
Db

QY 475 ATGAGTGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 533
479 ATGAGTGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
Db

QY 534 CCTCTGCTCAGCTCCCAAGTGTGGGACCAAGATGACACCTACCTGCTGCTA 593
539 CCTCTGCTCAGCTCCCAAGTGTGGGACCAAGATGACACCTACCTGCTGCTG--TA 597
Db

QY 594 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 653
598 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 656
Db

QY 654 GTGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 713
657 GTGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 716
Db

QY 714 TGGCCAGGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
717 TGGCCAGGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 776
Db

QY 774 GTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832
777 GTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836
Db

QY 833 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
837 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
Db

QY 892 CCGGCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 951
895 CCGGCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 954
Db

QY 952 GAGTCAATGAGCAATCTGCTCAGTCACTGCTGCTCCGGGCTCAGGATGCT 1011
955 GAGTCAATGAGCAATCTGCTCAGTCACTGCTGCTCCGGGCTCAGGATGCT 1012
Db

QY 1012 CCGTCTCAGGCTCCCAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1071
1013 CCGTCTCAGGCTCCCAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1070
Db

QY 1072 TTTGATTTTATGAGGCGGGGTTTACCATTTTGTAGGCT-GGTCTCAACTGCT 1130

Db 1071 TTTGATTTTCATGAGCGGG--TTTACATATTTCAGGCTGGGGTCTCAACTCT 1128
Qy 1131 GACCTCAGGTGACCCACTGCTGCTTCCAAAGTCTGGATTACAGCGGTGACCA 1190
Db 1129 GACCTCAGGTGACCCACTGCTGCTTCCAAAGTCTGGATTACAGCGGTGACCA 1188
Qy 1191 CCTCAGCGGCGGTATTTAGATAAATAATGTGCAATGGGGGTCTGTATGT 1250
Db 1189 CCTCAGCGGCGGTATTTAGATAAATAATGTGCAATGGGGGTCTGTATGT 1246
Qy 1251 TGCCAGGCTGCTCAACTTGTGCTTCAATGCAATCTTCCAAATGAGCCACAACAC 1310
Db 1247 TGCCAGGCTGCTCAACTTGTGCTTCAATGCAATCTTCCAAATGAGCCACAACAC 1306
Qy 1311 CAGCAGTACATTTTAAACAGTATCTTTATTTAGTATCTGAAAGTATACA 1370
Db 1307 CAGCAGTACATTTTAAACAGTATCTTTATTTAGTATCTGAAAGTATACA 1366
Qy 1371 ATAAACATGTCAAC 1385
Db 1367 ATAAACATGTCAAC 1381

RESULT 15
US-09-964-667-3
Sequence 3, Application US/09964667
Publication No. US20030033621A1
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
Mands, Jack R.
TITLE OF INVENTION: Transgenic Animals and Cell Lines for
Screening Drugs Effective for the Treatment or Prevention
of Alzheimer's Disease
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
STREET: 1100 New York Ave., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/964,667
FILING DATE: 28-Sep-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0609,4370000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1381 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: CDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-964-667-3

Query Match 74.9% Score 1080.2; DB 12; Length 1381;
Best Local Similarity 94.4%; Pred. No. 26e-304;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;
Qy 2 TTTTATTTTGAAGATGAGTTTCGCTCTGTGTCAGGCGAGGTGAGTCAATGGCGCAAT 61

Db 1 TTTTATTTTGAAGATGAGTTTCGCTCTGTGTCAGGCGAGGTGAGTCAATGGCGCAAT 60
Qy 62 CTCACCTCAGCGCAACCTCCGCTCCCGGTTCAAGCAATTCCTGCTCAGCTCCCC 121
Db 61 CTCACCTCAGCGCAACCTCCGCTCCCGGTTCAAGCAATTCCTGCTCAGCTCCCC 120
Qy 122 AGTACTGGGATTTACAGCATGTGACACCGCTGGCTAATTTGTAATTTTATTTT 181
Db 121 AGTACTGGGATTTACAGCATGTGACACCGCTGGCTAATTTGTAATTTTATTTT 179
Qy 182 AGAGATGAGATT--CTCCATGTTGTGAGGCTGCTGTAACCTCCACCTCAATATAT 239
Db 180 AGAGATGAGATTAACTCATATGTTGTCAGGCTGGTCTGTAACCTCCACCTCAATATAT 239
Qy 240 CCTCCGCTGCGCTCCCAAGTCT--AGATACAGACTGCGCCACCATGCGCGG--CT 295
Db 240 CTCCGCTGCGCTCCCAAGTCTGATTTACAGCATGAGCCACCATGCGCGGCT 299
Qy 296 CTGCTGCTAATTTTGTGTAAGACAGGTTTCACTGATG--TGCCCAAGCTGGTCTC 354
Db 300 CTGCTGCTAATTTTGTGTAAGACAGGTTTCACTGATGTTGCCAAGCTGGTCTC 359
Qy 355 CTGACCTCAGCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 414
Db 360 CTGACCTCAGCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418
Qy 415 CCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
Db 419 CCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 478
Qy 475 ATGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 533
Db 479 ATGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 538
Qy 534 CTTCTGCTCAGCTCCCAAGTACGTGGAGCAAGACATGACACACACACACACACAC 593
Db 539 CTTCTGCTCAGCTCCCAAGTACGTGGAGCAAGACATGACACACACACACACACAC 597
Qy 594 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 653
Db 598 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 656
Qy 654 GTGACGTGCGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 713
Db 657 GTGACGTGCGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
Qy 714 TGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 773
Db 717 TGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 776
Qy 774 GATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 832
Db 777 GATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 836
Qy 833 TTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
Db 837 TTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
Qy 892 CCCGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 951
Db 895 GCCGCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 954
Qy 952 GAGTCAATGCGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1011
Db 955 GAGTCAATGCG--CAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
Qy 1012 CTTGCTCAGCTCCCAAGCAGCTGAGGATTTAGGCGACCTGCGCACACACCGCTAAT 1071
Db 1013 CTTGCTCAGCTCCCAAGCAGCTGAGGATTTAGGCG--ACCTGACACACACCGCTAAT 1070
Qy 1072 TTTTATTTTCAATTTAGAGCGGGGTTTCAACATATTTTGTACAGCT--GGTCTCAACTCT 1130

Mon Jul 21 09:56:17 2003

us-09-380-203-1.rnpb

Page 16

Db	1071	TTTGTATTTTCCTATGAGAGGGG--TTTACCAATATTGTGACGGCTGGCTCTAAACTCT	1128
QY	1131	GACCTCAGGTGACCCACCTGCTCAGGCTCTTCCAAAGTGTGGATATTACAGSGGTGAGCCA	1190
Db	1129	GACCTCAGGTGACCCACCTGCTCAGGCTCTTCCAAAGTGTGGATATTACAGSGGTGAGCCA	1188
QY	1191	CTTCACCCAGCCGGCTAATTAGATAAAAAATATGTACGAATGGGGGGCTTGGTATCT	1250
Db	1189	CTTCACCCAGCCGGCTAATTAGATAAAAAATATGTACGAATGGGGG--TCTGGTATGT	1246
QY	1251	TGGCCAGCGTGGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACACC	1310
Db	1247	TGGCCAGCGTGGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACACC	1306
QY	1311	CAGCGACGTCAATTTTTTAAACAGTTACATCTTTATTTTGTACTAGCAAGATTAATACA	1370
Db	1307	CAGCGACGTCAATTTTTTAAACAGTTACATCTTTATTTTGTACTAGCAAGATTAATACA	1366
QY	1371	ATAAACATGTCAAC 1385	
Db	1367	ATAAACATGTCAAC 1381	

```
Search completed: July 20, 2003, 09:20:31
Job time : 341 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2003, 07:10:06 ; Search time 77 Seconds
(without alignments)
5743.219 Million cell updates/sec

Title: US-09-380-203-1

Perfect score: 1442

Sequence: 1 TTTTTCATGACATGACGAG.....TTAACAAAGCTTAGACCA 1442

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_NA:*
2: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1442	100.0	1442	2	US-08-454-557C-120
2	1442	100.0	1442	2	US-08-340-426D-120
3	1442	100.0	1442	2	US-08-450-673C-120
4	1223.4	84.8	1418	5	PCT-US95-17111A-120
5	1080.2	74.9	1381	2	US-08-454-557C-49
6	1080.2	74.9	1381	2	US-08-340-426D-49
7	1080.2	74.9	1381	2	US-08-450-673C-49
8	1080.2	74.9	1381	5	PCT-US95-17111A-49
9	475.4	33.0	14796	4	US-09-975-080-35
10	475.4	33.0	14796	4	US-09-630-706-10
11	475.4	33.0	14796	4	US-09-496-694B-3
12	416.4	28.9	59065	4	US-09-813-817-3
13	416.4	28.9	59065	4	US-09-978-197-3
14	412.2	28.6	43950	4	US-09-735-934A-3
15	400.8	27.8	5543	2	US-08-687-080-101
16	393.6	27.3	4421	4	US-08-257-963B-9
17	393.6	27.3	4421	4	US-08-367-841A-9
18	393.6	27.3	4421	4	US-08-520-373D-6
19	393.6	27.3	4421	5	PCT-US95-07201-9
20	393	27.3	53526	3	US-08-658-136-2
21	393	27.3	53577	3	US-08-658-136-1
22	392	27.2	62804	4	US-09-800-960-3
23	386.4	26.8	43950	4	US-09-735-934A-3
24	386	26.8	31571	1	US-08-323-443B-1
25	376.6	26.1	99500	4	US-09-798-096-10
26	374.6	26.0	14636	4	US-09-173-914-6
27	374.4	26.0	2598	4	US-09-026-033-18

28	371.4	25.8	26664	4	US-09-564-805-28	Sequence 28, Appl
29	370	25.7	4793	1	US-09-561-497-10	Sequence 10, Appl
30	367.4	25.5	6769	4	US-08-480-784-20	Sequence 20, Appl
31	367.4	25.5	6769	1	US-08-483-553-20	Sequence 20, Appl
32	367.4	25.5	6769	1	US-08-487-002-20	Sequence 20, Appl
33	367.4	25.5	6769	1	US-08-483-554B-20	Sequence 20, Appl
34	367.4	25.5	6769	4	US-08-488-011B-20	Sequence 20, Appl
35	367.4	25.5	6769	1	US-08-850-727-20	Sequence 20, Appl
36	367.4	25.5	6769	5	PCT-US95-10202-20	Sequence 20, Appl
37	367.4	25.5	6769	5	PCT-US95-10203-20	Sequence 20, Appl
38	367.4	25.5	6769	5	PCT-US95-10220-20	Sequence 20, Appl
39	357.2	24.8	8392	1	US-08-080-255-6	Sequence 6, Appl1
40	357.2	24.8	8392	3	US-08-465-713-6	Sequence 6, Appl1
41	357.2	24.8	8392	5	PCT-US93-05857-6	Sequence 6, Appl1
42	356.8	24.7	84495	4	US-09-797-906-3	Sequence 3, Appl1
43	356.6	24.7	45716	4	US-08-965-048-5	Sequence 5, Appl1
44	356.6	24.7	45989	4	US-08-965-048-6	Sequence 6, Appl1
45	355.8	24.7	3035	1	US-08-726-725-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-454-557C-120
Sequence 120, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Mandis, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ. ID NO.: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 1442 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 15..1139
US-08-454-557C-120
Query Match
Best Local Similarity 100.0%: Score 1442: DB 2: Length 1442:
Matches 1442: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY 1 TTTTTCATGACATGACGAGTTTCGCTCTGTCGCCAGCGTGCAGTGCATGCGCA 60

```
Db 1 TTTTGTGAGATGAGATTTTGTGCTTTGTTGCTCCAGGCTGAGGCAATGCGCCAA 60
QY 61 TCTGAGCTACCGGCAACCTCCGCGCTCCGCGGTTCAAGGATTTCTGCTCCAGCTCC 120
Db 61 TCTGAGCTACCGGCAACCTCCGCGCTCCGCGGTTCAAGGATTTCTGCTCCAGCTCC 120
QY 121 CAGTAGCTGGGATTTACAGCATGTGACACCGCTCCGCTAATTTTGTATTTTGTAG 180
Db 121 CAGTAGCTGGGATTTACAGCATGTGACACCGCTCCGCTAATTTTGTATTTTGTAG 180
QY 181 TAGAGATGAGATTTCTCATGTGTGACAGGCTGTGAACTCCGACCTCAGATGATC 240
Db 181 TAGAGATGAGATTTCTCATGTGTGACAGGCTGTGAACTCCGACCTCAGATGATC 240
QY 241 CCTCCGCTCCGCTCCCAAAAGTCTGATACAGAGATGCGACATGCGCGCTCGCC 300
Db 241 CCTCCGCTCCGCTCCCAAAAGTCTGATACAGAGATGCGACATGCGCGCTCGCC 300
QY 301 TGGCTAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGTCTCTGAGC 360
Db 301 TGGCTAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGTCTCTGAGC 360
QY 361 TCAAGCAGTCACCTGCTCAGCCTCCCAAGTGTGGGATTTACAGGCTGCAGCGCTGC 420
Db 361 TCAAGCAGTCACCTGCTCAGCCTCCCAAGTGTGGGATTTACAGGCTGCAGCGCTGC 420
QY 421 CTGGCCTTTTATTTTATTTTATTTTAAAGACAGGTGTCCCACTTTACCCAGATGAAG 480
Db 421 CTGGCCTTTTATTTTATTTTATTTTAAAGACAGGTGTCCCACTTTACCCAGATGAAG 480
QY 481 TGCAGTGTGTGATTCACAGCTCAGCTCAGCTTCAACTCTGATGATCAAGATCTCTCG 540
Db 481 TGCAGTGTGTGATTCACAGCTCAGCTCAGCTTCAACTCTGATGATCAAGATCTCTCG 540
QY 541 CCTCAGCTCCCAAGTGTGGGACCAAGACATGACACCTACACCTGAGATCTCTCTG 540
Db 541 CCTCAGCTCCCAAGTGTGGGACCAAGACATGACACCTACACCTGAGATCTCTCTG 540
QY 601 TTTTATTTTATTTTATTTTGTGAGACAGATGATCAACTCTGATGATCAAGATCTCTCG 660
Db 601 TTTTATTTTATTTTATTTTGTGAGACAGATGATCAACTCTGATGATCAAGATCTCTCG 660
QY 661 GCGCAGATTTGGCTCAGCTGCAACCTGCGCTCCGCGGTTCAAGTATTTCTGCGCCCA 720
Db 661 GCGCAGATTTGGCTCAGCTGCAACCTGCGCTCCGCGGTTCAAGTATTTCTGCGCCCA 720
QY 721 GCGCTGAGTGTGGGACTAGACAGCGCCCAACAGCTTATTTTGTATTTT 780
Db 721 GCGCTGAGTGTGGGACTAGACAGCGCCCAACAGCTTATTTTGTATTTT 780
QY 781 TAGTAGAGATGGGTTACCATGTTCGACAGTGTGATGATGATGAGACCTTGTATC 840
Db 781 TAGTAGAGATGGGTTACCATGTTCGACAGTGTGATGATGATGAGACCTTGTATC 840
QY 841 TGGCTGCGCTGGGCTCCCAAAAGTCTGAGTTACAGGCTGAGACCAACCGCGCTTA 900
Db 841 TGGCTGCGCTGGGCTCCCAAAAGTCTGAGTTACAGGCTGAGACCAACCGCGCTTA 900
QY 901 TTTTATTTTATTTTGTGTAATGATCACTGATGATGATGATGATGATGATGATGATG 960
Db 901 TTTTATTTTATTTTGTGTAATGATCACTGATGATGATGATGATGATGATGATGATG 960
QY 961 GGGCAATTCGCGCTCAGTCAACCTGCTCCGCGGTTCAAGGATTTCTCTGCTCA 1020
Db 961 GGGCAATTCGCGCTCAGTCAACCTGCTCCGCGGTTCAAGGATTTCTCTGCTCA 1020
QY 1021 GCGTCCCAACAGCTGGGATTTAGGGGACCTGCGACCAACCGCGCTATTTTGTATTT 1080
Db 1021 GCGTCCCAACAGCTGGGATTTAGGGGACCTGCGACCAACCGCGCTATTTTGTATTT 1080
QY 1081 TCATTAGAGCGGGGTTTACCATTTTGTAGGCTGTCTCAAACTCTGACCTCAGGT 1140
Db 1081 TCATTAGAGCGGGGTTTACCATTTTGTAGGCTGTCTCAAACTCTGACCTCAGGT 1140
```

```
Db 1081 TCATTAGAGCGGGGTTTACCATTTTGTAGGCTGTCTCAAACTCTGACCTCAGGT 1140
QY 1141 GACCCAGCTGCGCTCAGCTTCCAAAGTGTGGGATTTACAGGCTGAGGCAACCTCAGCT 1200
Db 1141 GACCCAGCTGCGCTCAGCTTCCAAAGTGTGGGATTTACAGGCTGAGGCAACCTCAGCT 1200
QY 1201 CCGGCTAATTTTGTATTAATAAATATGTAGCAATGGGGGCTTGTCTATGTTGCCAGCT 1260
Db 1201 CCGGCTAATTTTGTATTAATAAATATGTAGCAATGGGGGCTTGTCTATGTTGCCAGCT 1260
QY 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCCAAATGAGCCACACACCGCAGCTCA 1320
Db 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCCAAATGAGCCACACCGCAGCTCA 1320
QY 1321 CATTTTAAACAGTACATCTTATTTTGTATGATAGTAAGATTAATTAACATGT 1380
Db 1321 CATTTTAAACAGTACATCTTATTTTGTATGATAGTAAGATTAATTAACATGT 1380
QY 1381 CAACCTGCAAAATTCAGTAGTAACAGAGCTCTTTATTAACCTTTAAACAAAGCTTAGAG 1440
Db 1381 CAACCTGCAAAATTCAGTAGTAACAGAGCTCTTTATTAACCTTTAAACAAAGCTTAGAG 1440
QY 1441 CA 1442
Db 1441 CA 1442
```

RESULT 2

US-08-340-426D-120

Sequence 120, Application US/08340426D

Patent No. 5948634

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609, 3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2540

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 120:

SEQUENCE CHARACTERISTICS:

LENGTH: 1442 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 15..1139

US-08-340-426D-120

Query Match

100.0%; Score 1442; DB 2; Length 1442;

LOCATION: 15..1139.
US-08-450-673C-120

Query Match 100.0%; Score 1442; DB 2; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTTTGTGAGTGGAGTTTTCCTCTGTTGGCCAGAGCTGGAGTGCATGCGCGAA 60
Db 1 TTTTGTGAGTGGAGTTTTCCTCTGTTGGCCAGAGCTGGAGTGCATGCGCGAA 60
QY 61 TCTCAGCTACCCCAACTCCGCTCCGCGTTCAAGCATTCTCTGCTCAGCTCC 120
Db 61 TCTCAGCTACCCCAACTCCGCTCCGCGTTCAAGCATTCTCTGCTCAGCTCC 120
QY 121 CAGTACCTGGAGTTAAGGAGTGTGACACCAAGCTGGCTAATTTTGTATTTT 180
Db 121 CAGTACCTGGAGTTAAGGAGTGTGACACCAAGCTGGCTAATTTTGTATTTT 180
QY 181 TAGAGATGAGATTTCATGTTGGTCAAGCTGTGTCTGACACTCCGACCTG 240
Db 181 TAGAGATGAGATTTCATGTTGGTCAAGCTGTGTCTGACACTCCGACCTG 240
QY 241 CCGCTCTCGGCTCCCAAGTGTCTAGATACAGAGTGGCCACCATGCGCTCG 300
Db 241 CCGCTCTCGGCTCCCAAGTGTCTAGATACAGAGTGGCCACCATGCGCTCG 300
QY 301 TGGCTAATTTTGTGTAGAAACAGGTTTCACTGATGCGCCAGCTGTCTGTAG 360
Db 301 TGGCTAATTTTGTGTAGAAACAGGTTTCACTGATGCGCCAGCTGTCTGTAG 360
QY 361 TCAAGAGTCCACCTGCTCAGCTCCCAAGTGTCTGAGATTACAGGCTGAG 420
Db 361 TCAAGAGTCCACCTGCTCAGCTCCCAAGTGTCTGAGATTACAGGCTGAG 420
QY 421 CTGGCTTTTATTTTATTTTATTTTAAAGACAGGTGCTCCACTTTTACAG 480
Db 421 CTGGCTTTTATTTTATTTTATTTTAAAGACAGGTGCTCCACTTTTACAG 480
QY 481 TGCAGTGTGTGATACAGCTCAGCTTCAACTCTGAGATCAAGCATCTCTCT 540
Db 481 TGCAGTGTGTGATACAGCTCAGCTTCAACTCTGAGATCAAGCATCTCTCT 540
QY 541 CCTCAGCTCCCAAGTGTGAGACAGAGTGTCACTGTGACACCAAGCTGAG 600
Db 541 CCTCAGCTCCCAAGTGTGAGACAGAGTGTCACTGTGACACCAAGCTGAG 600
QY 601 TTTTATTTTATTTTATTTTATTTTAAAGACAGTGTCACTGTGACACCA 660
Db 601 TTTTATTTTATTTTATTTTATTTTAAAGACAGTGTCACTGTGACACCA 660
QY 661 GCGCAATCTGGCTCAGCTCAGCTGCTCCGCTCCGCGTTCAAGTATCTCT 720
Db 661 GCGCAATCTGGCTCAGCTCAGCTGCTCCGCTCCGCGTTCAAGTATCTCT 720
QY 721 GCTCTGAGTAGTGTGAGTACAGGCTCCAGCCTAGCTAATTTTGTATTT 780
Db 721 GCTCTGAGTAGTGTGAGTACAGGCTCCAGCCTAGCTAATTTTGTATTT 780
QY 781 TAGTAGAGATGGGTTTACCATGTTGCCAGTGTGATCTGTGACCTTGTG 840
Db 781 TAGTAGAGATGGGTTTACCATGTTGCCAGTGTGATCTGTGACCTTGTG 840
QY 841 TGCTGCTCGGCTCCCAAGTGTGAGTGTGAGGCTGAGCCACACCGCGCTTA 900
Db 841 TGCTGCTCGGCTCCCAAGTGTGAGTGTGAGGCTGAGCCACACCGCGCTTA 900
QY 901 TTTTATTTTATTTTGTGTTGAATGGAATCTACTCTGTACCGAGGCTGAG 960
Db 901 TTTTATTTTATTTTGTGTTGAATGGAATCTACTCTGTACCGAGGCTGAG 960
QY 961 GGGCAATCTGGCTCAGCTCAGCAACTCTGCTCCGCGCTCAAGCATTTCT 1020
Db 961 GGGCAATCTGGCTCAGCTCAGCAACTCTGCTCCGCGCTCAAGCATTTCT 1020
```

```
Db 961 GGGCAATCTGGCTCAGCTCAGCAACTCTGCTCCGCGCTCAAGCATTTCTCTCTCA 1020
QY 1021 GCGTCCCAAGCAGCTGGAGTTACGGGACCTGCCACACACCGCTAATTTTGTATTT 1080
Db 1021 GCGTCCCAAGCAGCTGGAGTTACGGGACCTGCCACACACCGCTAATTTTGTATTT 1080
QY 1081 TCATTAGAGAGGGGTTTACCATATTGTGAGGCTGTCTCAACTCTGAGCTCAGGT 1140
Db 1081 TCATTAGAGAGGGGTTTACCATATTGTGAGGCTGTCTCAACTCTGAGCTCAGGT 1140
QY 1141 GACCCAGCTGCTCAGCTTCAAGTGTGAGATTACAGGCTGAGCCACCTCAGCCAG 1200
Db 1141 GACCCAGCTGCTCAGCTTCAAGTGTGAGATTACAGGCTGAGCCACCTCAGCCAG 1200
QY 1201 CCGGCTAATTTAGTAAATTTAGTAAATGAGCAATGCGGGGCTTGTATGTGCGCAG 1260
Db 1201 CCGGCTAATTTAGTAAATTTAGTAAATGAGCAATGCGGGGCTTGTATGTGCGCAG 1260
QY 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCCAAATGAGCCACACCCAGCTA 1320
Db 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCCAAATGAGCCACACCCAGCTA 1320
QY 1321 CATTTTAAACAGTTACATCTTATTTTGTATCTAGAAAGTAAATACATGAT 1380
Db 1321 CATTTTAAACAGTTACATCTTATTTTGTATCTAGAAAGTAAATACATGAT 1380
QY 1381 CAAACTGCAAAATCTAGTAAACAGAGTCTTTTAACTTTTAAACAAGCTTTAGAG 1440
Db 1381 CAAACTGCAAAATCTAGTAAACAGAGTCTTTTAACTTTTAAACAAGCTTTAGAG 1440
QY 1441 CA 1442
Db 1441 CA 1442

RESULT 4
PCT-US95-17111A-120
Sequence 120, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Wand's, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 1418 base pairs
```


TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: 14..1418
PCT-US95-17111A-120

Query Match 84.8%; Score 1223.4; DB 5; Length 1418;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1375; Conservative 0; Mismatches 31; Indels 13; Gaps 12;

```

OY 2 TTTTGTGAGATGAGATTTGCTCTTGTGCGCAGGCTGAGTCAATGGGCAAT 61
DB 1 TTTTGTGAGATGAGATTTGCTCTTGTGCGCAGGCTGAGTCAATGGGCAAT 60
OY 62 CTGAGCTACCGCAACCTCCGCTCCGGGTTCAAGCATTCCTGGCTCAGCTCCG 121
DB 61 CTGAGCTACCGCAACCTCCGCTCCGGGTTCAAGCATTCCTGGCTCAGCTCCG 120
OY 122 AGTA-GCTGGATTTACAGCATGTGACCCAGCGCTGCTAATTTGTATTTTGTAG 180
DB 121 AGTAGGCTGGGATTTACAGCATGTGCA-CACGCTGGCTAATTTGTATTTTGTAG 179
OY 181 TAGAGATGAGATTTCTCCATGTTGGTCAGGCTGTCTGAACTCCGACCTGATGATC 240
DB 180 TAGAGATGAGATTTCTCCATGTTGGTCAGGCTGTCTGAACT-CCGACCTGATGATC 238
OY 241 CTTCCGCTTCGGGCTCCCAAGTGTATATACAGGACTGGCCACCATGCCGG-CTGTCG 299
DB 239 CTCCGCTTCGGGCTCCCAAGTGTATATACAGGACTGGCCACCATGCCGGCTGTCG 298
OY 300 CTGGCTAATTTTGTGTAGAAACAGGGTTTCAGTGTGCGCAAGCTGGCTCCTGAG 359
DB 299 CTGGCTAATTTTGTGTAGAAACAGGGTTTCAGTGTGCGCAAGCTGGCTCCTGAG 358
OY 360 CTCAAGCAGTCACCTGCTCAGCTCCCAAGTGTGAGATTACAGGCTGACACCGTG 419
DB 359 CTCAAGCAGTCACCTGCTCAGCTCCCAAGTGTGAGATTACAGGCTGACACCGTG 418
OY 420 CTTGGCTTTTATTTTATTTTATTTTAAAGACAGAGTGTCCGACTTTACAGAGATGAA 479
DB 419 CTTGGCTTTTATTTTATTTTATTTTAAAGACAGAGTGTCCGACTTTACAGAGATGAA 478
OY 480 GTGAGCTGTGATCAGCTCAGCTGAGCTTCACTCCGATGAGATGAGATGATCTCT 539
DB 479 GTGAGCTGTGATCAGCTCAGCTGAGCTTCACT-CTGAGATGAGATGATCTCTCT 537
OY 540 GCTTCAGGCTCCC-AAAGTGTGGGACCAAGACATGACCACTACCTGGCTAATTT 598
DB 538 GCTTCAGGCTCCC-AAAGTGTGGGACCAAGACATGACCACTACCTGGCTAATTT 597
OY 599 TATTTTATTTTATTTTATTTTGAAGACAGATCTCACTCTGACCCAGGCTGGAGTGA 658
DB 598 TATTTTATTTTATTTTATTTTGAAGACAGATCTCACTCTGACCCAGGCTGGAGTGA 657
OY 659 GTGGGCAATCTTGCTGCTACTGCAACCTCTGCTCCGGGTTCAAGTATTTCTCTGCC 718
DB 658 GTGGGCAATCTTGCTGCTACTGCAACCTCTGCTCCGGGTTCAAGTATTTCTCTGCC 717
OY 719 CAGCTCTCTGATAGCTGGGACTACAGGCGCCACACAGCTAGTAAATTTTGTATTT 778
DB 718 CAGCTCTCTGATAGCTGGGACTACAGGCGCCACACAGCTAGTAAATTTTGTATTT 777
OY 779 TTTAGTAAAGATGGGG-TTCACATGTTCCGACAGTTAT-CTTGATCTCTGGACCTTGT 836
DB 778 TTTAGTAAAGATGGGGTTTACACATGTTCCGACAGTTATGCTTGTGACCTTGT 837
OY 837 GATCTGCTGCTGGGCTCCCAAGTGTGGGATTTACAGG-CGTGAGCCACACAGCCGG 895
DB 838 GATCTGCTGCTGGGCTCCCAAGTGTGGGATTTACAGGCGTGGAGCCACCGCCGG 897
OY 896 GCTTATTTTATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGAGT 955

```

```

DB 898 GCTTATTTTATTTTGTGTTTGTGGAATGGAATCTCACTCTGTACCAGGCTGAGT 957
OY 956 GCATGGCCAAATCTCGGCTCAGTCAACCTCTGCTCCCGGCTCAAGCATTTCTCG 1015
DB 958 GCATGGCCAAATCTCGGCTCAGTCAACCTCTGCTCCCGGCTCAAGCATTTCTCG 1017
OY 1016 TCTCAGCTCCCAACACAGCTGGGATTTACGGGACCTGCAACACCCCGTAAATTTTG 1075
DB 1018 TCTCAGCTCCCAACACAGCTGGGATTTACGGGACCTG-CACCAACCCCGTAAATTTTG 1076
OY 1076 TATTTTCAATTAGAGCGGGGTTTCCACATATTTGTGAGGCTGTCTCAACTCTGAGCT 1135
DB 1077 TATTTTCAATTAGAGCGGGGTTTCCACATATTTGTGACGGCTGTCTCAACTCTGAGCT 1136
OY 1136 CAGGTGACCCACCTGCTCAGCTTCCAAAGTGTGGGATTTACAGGCTGAGCCATCA 1195
DB 1137 CAGGTGACCCACCTGCTCAGCTTCCAAAGTGTGGGATTTACAGGCTGGA--CGCTCA 1194
OY 1196 CCCAGCGGCTAATTTAGATTAATAAATATGAGCAATGGGGGCTTGTATTTGGCC 1255
DB 1195 CCCAGCGGCTAATTTAGATTAATAAATATGAGCAATGGGGGCTTGTATTTGGCC 1254
OY 1256 AGGCTGTCTCAACTCTGCTTCATGCAATCTTCCAAATGAGCCACACCCAGCC 1315
DB 1255 AGGCTGTCTCAACTCTGCTTCATGCAATCTTCCAAATGAGCCACACCCAGCC 1314
OY 1316 AGTCATTTTAAACAGTTACATCTTATTTAGTATAGTAAAGTAAATACATGATAA 1375
DB 1315 AGTCACA-TTTTAAACAGTTACATCTTATTTAGTATAGTAAAGTAAATACATGATAA 1373
OY 1376 CATGTCAAACCTGCAATTTCACTAGTACAGAGTCTTT 1414
DB 1374 ATGGCGGAACCTGCAATTTCACTAGTACAGAGTCTTT 1412

```

RESULT 5

US-08-454-557C-49
Sequence 49, Application US/0845457C

Patent No. 5830670

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Mands, Jack R.

TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/454,557C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609,3840003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 1381 base pairs

TYPE: nucleic acid

Query Match 74.9%; Score 1080.2; DB 2; Length 1381;
 Best Local Similarity 94.4%; Pred. No. 3.5e-292;
 Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

```

OY 2 TTTTATTTTATGAGATGAGTTTGGCTCTGTGGCCAGGCTGAGTCGAATGGCGCAAT
   |||||||
Db 1 TTTTATTTTATGAGATGAGTTTGGCTCTGTGGCCAGGCTGAGTCGAATGGCGCAAT
   |||||||
OY 62 CTCAGCTACGGCAACCTCCGGGTTCAAGGATTCCTGCTCAGCCTCCGCC
   |||||||
Db 61 CTCAGCTACGGCAACCTCCGGGTTCAAGGATTCCTGCTCAGCCTCCGCC
   |||||||
OY 122 AGTAGCTGGATTAACAGCATGTGACCCAGCTCGCTAATTTTGTATTTTATAGT
   |||||||
Db 121 AGTAGCTGGATTAACAGCATGTGACCCAGCTCGCTAATTTTGTATTTTATAGT
   |||||||
OY 182 AGAGATGAGATT--CTCCATGTTGGTCAAGCTGTGCAACTCCGACCTCAGATGAT
   |||||||
Db 180 AGAGATGAGATTAACTCATGTTGGTCAGGCTGTGCAACTCCGACCTCAGATGAT
   |||||||
OY 240 CCTCCGCTGGGCTCCCAAGTGC---AGATACAGAGCTGGCCAGCATGCCGG--CT
   |||||||
Db 240 CTCCGCTGGGCTCCCAAGTGC---AGATACAGAGCTGGCCAGCATGCCGG--CT
   |||||||
OY 296 CTGCTGCTAATTTTGTGTAGAAACAGAGGTTTCACTGATG--TGCCCAAGCTGTCTC
   |||||||
Db 300 CTGCTGCTAATTTTGTGTAGAAACAGAGGTTTCACTGATGTTGCCCAAGCTGTCTC
   |||||||
OY 355 CTGAGCTAAGCAGTCCACCTGCTCAGCCTCCCAAGTGTGGGATTAACAGGCTGCAG
   |||||||
Db 360 CTGAGCTAAGCAGTCCACCTGCTCAGCCTCCCAAGTGTGGGATTAACAGGCTGCAG
   |||||||
OY 415 CCGTCCCTGGGCTTTTATTTTATTTTATTTTAAAGACAGAGGTCCTTACCCAGG
   |||||||
Db 419 CCGTCCCTGGGCTTTTATTTTATTTTATTTTAAAGACAGAGGTCCTTACCCAGG
   |||||||
OY 475 ATGAAGTCAAGTGTGTATGATCAGACCTCAGCAGCTTCAACTCTGATGATCAAGC--AT
   |||||||
Db 479 ATGAAGTCAAGTGTGTATGATCAGACCTCAGCAGCTTCAACTCTGATGATCAAGC--AT
   |||||||
OY 534 CCTCTGCTCAGCCTCCCAAGTGTGGGACCAAGACATGACACTACACCTGGCTA
   |||||||
Db 539 CCTCTGCTCAGCCTCCCAAGTGTGGGACCAAGACATGACACTACACCTGG--TA
   |||||||
OY 594 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
   |||||||
Db 598 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
   |||||||
OY 654 GTGCAAGTGGCAATCTTGGCTCACTGCAACCTGTGCTCCGGGTTCAAGTTATTTCTC
   |||||||
Db 657 GTGCAAGTGGCAATCTTGGCTCACTGCAACCTGTGCTCCGGGTTCAAGTTATTTCTC
   |||||||
OY 714 TGCCCAAGCTTCCTGAGTGTGGGACACAGGCGCCACACAGGCTAAGTTATTTT
   |||||||
Db 717 TGCCCAAGCTTCCTGAGTGTGGGACACAGGCGCCACACAGGCTAAGTTATTTT
   |||||||
OY 774 GTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
   |||||||
Db 777 GTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
   |||||||
OY 833 TTGTGATCTGCTGCTGCTGCT--CCCAAGTGTGGGATTAACAGCTGAGCCACAG
   |||||||
Db 837 TTGTGATCTGCTGCTGCTGCTGCT--CCCAAGTGTGGGATTAACAG--GTCTGATCTCAG
   |||||||
OY 892 CCGGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
   |||||||
Db 895 GCGGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
   |||||||
OY 952 GAGTCAATGGCAATCTGGGCTCACTGCAACCTGTGCTCCGGGTTCAAGGATTTCT
   |||||||
Db 955 GAGTCAATGG--CAATCTGGGCTACTGCAACCTGTGCTCCGGG--TCAAGGATTTCT
   |||||||
OY 1012 CCTGTCTAGGCTCCCAAGCAGCTGGGATTAAGGGGACCTGTCCACACACCCCGCTAATT
   |||||||
  
```

```

Db 1013 CCTGTCTAGGCTCCCAAGCAGCTGGGATTAAGGG--ACCTGACACACACCCCGCTAATT
   |||||||
OY 1072 TTTGATTTTATTTATGAGAGCGGGGTTTACCATATTTTGTACAGCT--GGTCTCAACTCT
   |||||||
Db 1071 TTTGATTTTATTTATGAGAGCGGG--TTTACCATATTTTGTACAGGCTGTCTCAACTCT
   |||||||
OY 1131 GACCTCAGGTGACCCACCTGCTCAGCCTTCCAAAGTCTGGGATTTACAGCGTGAGCCA
   |||||||
Db 1129 GACCTCAGGTGACCCACCTGCTCAGCCTTCCAAAGTCTGGGATTTACAGCGTGAGCCA
   |||||||
OY 1191 CCTCAGCAGCGGCTAATTTAGATTAATAATATGATGACATGGGGGCTTGGCTATGT
   |||||||
Db 1189 CCTCAGCAGCGGCTAATTTAGATTAATAATATGATGACATGGGGG--TCCTGATGT
   |||||||
OY 1251 TGCCAGGCTGTCTCAAACTTGTGGCTTATGATGATCTCTTCAATATGACACACACC
   |||||||
Db 1247 TGCCAGGCTGTCTCAAACTTGTGGCTTATGATGATCTCTTCAATATGACACACACC
   |||||||
OY 1311 CAGCAGTCAATTTTAAACAGTTACATCTTTTATTTAGATTAAGAAATATACA
   |||||||
Db 1307 CAGCAGTCAATTTTAAACAGTTACATCTTTTATTTAGATTAAGAAATATACA
   |||||||
OY 1371 ATAAACATGTCAAC 1385
   |||||||
Db 1367 ATAAACATGTCAAC 1381
   |||||||
  
```

RESULT 7
 US-08-450-673C-49
 ; Sequence 49, Application US/08450673C
 ; Patent No. 594888
 ; GENERAL INFORMATION:
 ; APPLICANT: de la Monte, Suzanne
 ; APPLICANT: Wands, Jack R.
 ; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
 ; TITLE OF INVENTION: of Alzheimer's Disease
 ; NUMBER OF SEQUENCES: 121
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005-3934
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/450,673C
 ; FILING DATE: 30-MAY-1995
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ludwig, Steven R.
 ; REGISTRATION NUMBER: 36,203
 ; REFERENCE/DOCKET NUMBER: 0609,3840004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 49:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1381 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; US-08-450-673C-49

Query Match 74.9%; Score 1080.2; DB 2; Length 1381;
 Best Local Similarity 94.4%; Pred. No. 3.5e-292;
 Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

```

OY 2 TTTTGTGAGATGAGATTTTGGCTCTGTTGGCCAGGCTGAGTGCATATGGGCAAT 61
DB 1 TTTTGTGAGATGAGATTTTGGCTCTGTTGGCCAGGCTGAGTGCATATGGGCAAT 60
OY 62 CTCAGCTACCCCAACCTCCGCTCCGCGGTCAAGCAATTCCTGCTCAGCCCTCC 121
DB 61 CTCAGCTACCCCAACCTCCGCTCCGCGGTCAAGCAATTCCTGCTCAGCCCTCC 120
OY 122 AGTAGCTGAGATTAAGGATGTGCACCCAGCTGGCTAAATTTGTTATTTTATG 181
DB 121 AGTAGCTGAGATTAAGGATGTGCACCCAGCTGGCTAAATTTGTTATTTTATG 179
OY 182 AGAGATGAGATT--CTCCATGTTGTGAGAGCTGTCTCAACTCCCACTCAGATAT 239
DB 180 AGAGATGAGATTAACTCATGTGTGAGAGCTGTCTCAACTCCCACTCAGATAT 239
OY 240 CCGTCCGCTCGGCTCCCAAGTGT--AGATACAGAGCTGGCCACCATGCGCGG--CT 295
DB 240 CCGTCCGCTCGGCTCCCAAGTGTCTAGATTAAGGATGAGAGCTGGCCACCATGCGCGGCT 299
OY 296 CTGCTGCTAATTTTGTGAGAAAGGTTTCACTGATG--TGCCCAAGCTGTCTC 354
DB 300 CTGCTGCTAATTTTGTGAGAAAGGTTTCACTGATGTTGCCAAGCTGTCTC 359
OY 355 CTGAGCTAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGAGATTACAGGCTGAG 414
DB 360 CTGAGCTAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGAGATTACAGGCTGAG 418
OY 415 CCGTCCGCTGCTTTTATTTTATTTTAAAGCAAGTGTCCCTCTTACCCAGG 474
DB 419 CCGTCCGCTGCTTTTATTTTATTTTAAAGCAAGTGTCCCTCTTACCCAGG 478
OY 475 ATGAAGTCAAGTGTGTGATCAGAGCTCAGAGCTTCACTCTGAGATCAAGC--AT 533
DB 479 ATGAAGTCAAGTGTGTGATCAGAGCTCAGAGCTTCACTCTGAGATCAAGC--AT 538
OY 534 CCGTCCGCTGAGCTCCCAAGTGTGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCA 593
DB 539 CCGTCCGCTGAGCTCCCAAGTGTGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCA 597
OY 594 ATTTTATTTTATTTTATTTTATTTTGAAGAGAGTCTCACTCTGAGAGCTGAG 653
DB 598 ATTTTATTTTATTTTATTTTATTTTGAAGAGTCTC--ACTCTGAGAGCTGAG 656
OY 654 GTGAGTGGGCAATTTGCTCAGTGCACCTGCTCCGCGGTCAAGTTATTC 713
DB 657 GTGAGTGGGCAATTTGCTCAGTGCACCTGCTCCGCGGTCAAGTTATTC 716
OY 714 TGCCCAAGCTCTGAGTGTGAGTCAAGGCGCCACAGCCCTAGCTAATTTT 773
DB 717 TGCCCAAGCTCTGAGTGTGAGTCAAGGCGCCACAGCCCTAGCTAATTTT 776
OY 774 GTATTTTATGAGATGGG--TTACACATGTTGCCAGGTTGATCTGTGAGC 832
DB 777 GTATTTTATGAGATGGGTTTACACATGTTGCCAGGTTGATCTGTGAGC 836
OY 833 TTGAGATCTGCTCCCTGCGCT--CCCAAGTGTGAGTTACAGGCTGAGCCACAG 891
DB 837 TTGAGATCTGCTCCCTGCGCTACCCAAAGTGTGAGTTACAG--GTGCTGAGCTCAC 894
OY 892 CCGGCTAATTTTATTTTGTGTTGAATGSAATCTCACTGTACCCAGCTG 951
DB 895 CCGGCTAATTTTATTTTGTGTTGAATGSAATCTCACTGTGTTACCCAGCTG 954
OY 952 GAGTGCATGGCAATCTGCTCAGTCAACCTGCTCCGCGCTCAAGGATCT 1011
DB 955 GAGTGCATGG--CAATCTGCTGCTACCTGCTCCGCGG--TCAAGGATCT 1012
OY 1012 CCGTCCAGCTCCCAAGAGCTGAGTCAAGGCACTGCGCACACACCCGCTAAT 1071
DB 1013 CCGTCCAGCTCCCAAGAGCTGAGTCAAGG--ACCTGAGCACACCCGCTAAT 1070
OY 1072 TTTGATATTTTATGAGGCGGTTTACCATATTTTGTAGCT--GCTGTCAACACTCT 1130

```

```

DB 1071 TTTGATATTTATGAGCGGG--TTTACATATTTTGTAGGCTGCTCAAACTCT 1128
OY 1131 GAGCTCAGTACCCACCTGCTCAGCTTCCAAAGTGTGGGATTTACAGGCTGAGCCA 1190
DB 1129 GAGCTCAGTACCCACCTGCTCAGCTTCCAAAGTGTGGGATTTACAGGCTGAGCCA 1188
OY 1191 CCTCAGCCAGCGGCTAATTTAGATAAAAAATGTAGCAATGGGGGCTTGTATGT 1250
DB 1189 CCTCAGCCAGCGGCTAATTTGAAATAAAATGTAGCAATGGGG--TGTGTATGT 1246
OY 1251 TGCCAGGCTGTCTCAACTCTGCTTCAATGCAATCTCCAAATGAGCCACACACC 1310
DB 1247 TGCCAGGCTGTCTCAACTCTGCTTCAATGCAATCTCCAAATGAGCCACACACC 1306
OY 1311 CAGCAGTACATTTTAAACAGTACATCTTATTTAGTACTAGAAAGTATACA 1370
DB 1307 CAGCAGTACATTTTAAACAGTACATCTTATTTAGTACTAGAAAGTATACA 1366
OY 1371 ATAAACATGTCAAC 1385
DB 1367 ATAAACATGTCAAC 1381

```

```

RESULT 8
PCT-US95-17111A-49
; Sequence 49, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
PCT-US95-17111A-49

```

```

Query Match 74.9%; Score 1080.2; DB 5; Length 1381;
Best Local Similarity 94.4%; Pred. No. 3,5e-292;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;
OY 2 TTTTGTGAGATGAGATTTTGGCTCTGTTGGCCAGGCTGAGTGCATATGGGCAAT 61

```

Db 1 TTTTGTGAGAGGAGTTTTCGCTGTTGCCCCAGGCTGAGATGCAATGGCGCAAT 60
QY 62 CTCAGCTCAGCCGCAACCTCCGCTCCCGGGTTCAGGATTCCTGCTCAGCTCCCC 121
Db 61 CTCAGCTCAGCCGCAACCTCCGCTCCCGGGTTCAGGATTCCTGCTCAGCTCCCC 120
QY 122 ACTAGCTGGGATTCAGGAGATGTCACCCAGCTCGGCTAATTTTGTATTTTGTAT 181
Db 121 ACTAGCTGGGATTCAGGAGATGTCACCCAGCTCGGCTAATTTTGTATTTTGTAT 179
QY 182 AAGAGATGAGTTT - CTCATGTTGTCAGGCTGCTGCTGAGTCTCCGAGCTAGAT 239
Db 180 AAGAGATGAGTTTAACTGATGTTGGTCAGGCTGCTGAGTCTCCGAGCTCAGAT 239
QY 240 CCTCCGCTCGGCTCCCAAGTGT - AGATACAGGAGCTGGCCACCATGGCCG - CT 295
Db 240 CTCCGCTCGGCTCCCAAGTGTGAGATTCAGGAGATGAGCCATGCTCCGCT 299
QY 296 CTGCTGCTAATTTTGTGTAAGAAACAGGGTTTCACTGATG - TGCCAAAGCTGCTC 354
Db 300 CTGCTGCTAATTTTGTGTAAGAAACAGGGTTTCACTGATGTTGCCAAGCTGCTC 359
QY 355 CTGAGCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
Db 360 CTGAGCTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418
QY 415 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474
Db 419 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478
QY 475 ATGAGTGCAGTGTGTCATCAGAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 533
Db 479 ATGAGTGCAGTGTGTCATCAGAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
QY 534 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 593
Db 539 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
QY 594 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 653
Db 598 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 656
QY 654 GTGAGTGGCGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 713
Db 657 GTGAGTGGCGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
QY 714 TGCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 773
Db 717 TGCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
QY 774 GTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 832
Db 777 GTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 836
QY 833 TTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
Db 837 TTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
QY 892 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951
Db 895 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 952 GAGTCAATGAGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1011
Db 955 GAGTCAATGAG - CAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
QY 1012 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1071
Db 1013 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1070
QY 1072 TTTGTATTTTCAATTAGAGGCGGGTTTCAACATATTGTCAGGCT - GGTCTCAACTCCT 1130
Db 1071 TTTGTATTTTCAATTAGAGGCGGG - TTTACCAATTTTGTGAGGCTGCTCAACTCCT 1128

QY 1131 GACCTCAGTGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1190
Db 1129 GACCTCAGTGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1188
QY 1191 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1250
Db 1189 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
QY 1251 TGCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1310
Db 1247 TGCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1306
QY 1311 CAGCCAGTCAATTTTAAAGAGTATGATGATGATGATGATGATGATGATGAT 1370
Db 1307 CAGCCAGTCAATTTTAAAGAGTATGATGATGATGATGATGATGATGATGAT 1366
QY 1371 ATAAAGTGTCAAC 1385
Db 1367 ATAAAGTGTCAAC 1381

RESULT 9

US-08-975-080-35
Sequence 35, Application US/08975080
Patent No. 6245523
GENERAL INFORMATION:
APPLICANT: Altieri, Dario C.
TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,080
FILING DATE: 20-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,435
FILING DATE: 20-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7176
TELEFAX: 202-467-7000
INFORMATION FOR SEQ. ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 14796 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-975-080-35

Query Match 33.0%; Score 475.4; DB 4; Length 14796;
Best Local Similarity 69.8%; Pred. No. 7.8e-123;
Matches 877; Conservative 0; Mismatches 321; Indels 59; Gaps 15;

QY 1 TTTTGTGAGAGGAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 5380 TTTTGTGAGATGAG - TTTCACTCTGTTGCTGCTGCTGCTGCTGCTGCTGCT 5438

[illegible]

Db 6500 TGTGTGTCCAGGCTGTGTCCTTGAACCTCTGAGCTCACGGGATCTTTTTCCTCGGGGCTCC 6550
 QY 1162 CAAAGTGTCTGGGATTTACAGGCGGTGAGCCACCTCACCGCGGCTAATTTAGATATAA 1218
 Db 6560 CAAAGTGTGAGATTTACAGGCGATGAGCCACCGTCTCCAGCTAGAGATTCATTTTAA 6616

RESULT 10
US-09-630-706-10
; Sequence 10, Application US/09630706
Date: 09/27/2009

; PATENT NO. 6277640
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett
 ;

```

; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF HER-3 EXPRESSION
; FILE REFERENCE: RTS-0053

```

; CURRENT APPLICATION NUMBER: US/09/630,706
 ; CURRENT FILING DATE: 2000-08-01
 ; NUMBER OF SEQ ID NOS: 94

```

; NUMBER OF SEQ ID NOS: 3
; SEQ ID NO 10
; LENGTH: 14796
;

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

NAME/KEY: CDS
LOCATION: (2811
NAME/KEY: CDS

```

; NAME/KEY: CDS
; LOCATION: (3174)...(3283)
; NAME/KEY: CDS

```

```

; LOCATION: (5158)..(5275)
; NAME/KEY: CDS
; LOCATION: (11955)..(12044)

```

US-09-630-706-10

Query Match	33.0%;	Score 475.4;	DB 4;	Length 14796;
Best Local Similarity	69.8%;	Pred. No. 7.8e-123;		
Matches 877;	Conservative 0;	Mismatches 321;	Indels 59;	Gaps 15;

Matches 877; Conservative 0; Mismatches 321; Indels 59; Gaps 15;

OY
1 TTTTCTTTTGAGATGAGATTTCGCTCTGTGTGCCACGACTGGAGTGCATGGCCCA 60

Db 5380 TTTTNTTTTTCGAGATGAG-TTTCACCTTTGTTTCCCAAGCTGGAGTCGATGGTGCAG 5438
 Oy 61 TCTACGCTACCCGCAACCTCCGCGCTCCCGGGTTCAAGGAGATTCTCTCCGCTTACGCTTCTTC 120

D_b 5439 TCTTGGCTCACTGCACCTCTGCCTCGGGTTCAAGTGAATTCTCTGCGCCATGACCTGCC 54198

QY 121 CAGTAGCTGGGATTACAGGCATGTGCACCCAGCTCGGTAATTTTGTATTTTCTTTAG 180

DB 5499 AAGTAACGTGGGATTAAAGGGAAGTGCACACACCAGCTAATTTT --TGTAATTTTAAAG 5555

Db 5556 TAGAGATGGGGTTTCACACACATGGCCCGAGGCTGCTTGAACTCCGAGACCTC--GTGATT 5613

[illegible]

Dy 301 TGCGTAAATTTTGTGCTAGAAACAGAGGTTTCACTGATGTGCCCAAGCTGGTCCCTG--- 357
Db 5614 CGCCCAACCTTGGCCCTCCCAAGTGTGTG---GATTACAGGCGGTGAACCCACCAAGCCCTGGC 5670

D_b

5671 TTTTCTTTTGTCCT-GAGACACAGATTTCACCTGTATCCCAAGCCTGAGATAGGGTGG 5729

QY 358 --AGCTCAGCAGTCCACCTGCCTCAGCCTCCCAAGTGTGGATTACAGCGGTGCAGC 415

Db 5730 CCTGATCTCGGATCATCTGCACACCTCCGCTCCTGGGCTCAAGTAGATTGGCTGCTTACG 5789

Db 5790 CTCCAGTAGCCGAGATTACAGGCATGTGCCACCACACACCAGGTAATTTTGTATTTT 5849

476 TGAAGTCACGTGGTGTGATCA--CAGCTCACTGCAGCCTTCAACTCCTGAGATCAAC-CA 532


```

: CURRENT APPLICATION NUMBER: US/09/978,197
: CURRENT FILING DATE: 2001-10-17
: PRIOR APPLICATION NUMBER: 09/813,817
: PRIOR FILING DATE: 2001-03-22
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 59065
: TYPE: DNA
: ORGANISM: Human
US-09-978-197-3

Query Match      28.9%; Score 416.4; DB 4; Length 59065;
Best Local Similarity 67.7%; Pred. No. 4.3e-106;
Matches 815; Conservative 0; Mismatches 326; Indels 63; Gaps 14;

QY 42 CTGGAGTCAGTAATGGGCGCAATCTCAGCTCAGCCGCAACCTCCGCTCCCGGGTTCAAGCAT 101
DB ||||| ||| ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 15163 CTGGAGGCGAGTGGCGCATCTCAGCAGACATGCAACCTCCATCTCCAGGTTCAAGCAT 15104
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 102 TCTCCTGCTGAGCTCCCGCAGTACGTGGATTTACAGCATGTGACCCAGCCTGCGCTA 161
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 15103 TCTCCTGCTGAGTCTCTAGTACTGATGATGAGGTGCTGCTGCTGCTGCTGCTGCTA 15044
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 162 ATTTTGTATTTTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 221
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 15043 ATTTT---TGTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 14987
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 222 CTCCGAGCTCAGATGATCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 280
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14986 CTCCGAGCTCAGATGATCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 14927
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 281 CCACCATCCCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 340
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14926 AGCCAGCGCGCCAGCATCTCAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 14867
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 341 CCCAAGCTG-----GTCCTGAGCTCA 363
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14866 CCCAAGCTGAGATGAGTGGCGCATCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 14807
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 364 AGCAGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14806 CGCATTCTCCGCTCAGCTCAGCTGAGTACGAGGAGTACGAGGCTGCTGCTGCTGCTGCT 14747
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 422 TGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 481
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14746 TGGCTAATTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 14687
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 482 CGAGTGTGTGATCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 541
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14866 -CGATCTCTGACCTTGTGATCCGCCGCCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 14628
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 542 CTCAGCCCTCCAGATGAGTGGAGCAAGA-----CATGCAACCTACACCTGCTGCTGCT 596
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14627 GTGAGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14568
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 597 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 656
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14567 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 14509
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 657 CAGTGGCCCAATCTTGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 716
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14508 CAGTGGCCCAATCTTGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 14449
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 717 CCCAGCTCCG---TGAGTAGCTGGGAGTACAGGGCC---CCAGCAGGCTGAGCTAATTT 770
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14448 CTCAGCTCCGAGTAGTGGAGTACAGGGAGTACAGGGAGTACAGGGAGTACAGGGAGTACAG 14391
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 771 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 829
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 14390 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 14331
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 830 ACCT--TGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

DB 14330 ACCTCAAGTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 14271
QY 888 CAGCCGCGGCTATTT--TTAATTTTGTGTTGTTGAATGGAATCTCAGCTGTTACCA 946
DB 14270 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14211
QY 947 GCGTGGAGTGAATGGCAAAATCTGCGCTCAGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCT 1006
DB 14210 GCGTGGAGTGAATGGCAAAATCTGCGCTCAGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCT 14151
QY 1007 ATTTCTGCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1066
DB 14150 ATTTCTGCTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 14091
QY 1067 TAAATTTTGTATTTTATTTATTAAGGGGGGTTTACCATATTTTGTAGGCTGCTC--TCAAA 1125
DB 14090 TAAATTTTGTATTTTATTTTGTGGGATGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 14031
QY 1126 CTCTGAGCTCAGTGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185
DB 14030 CTCTGAGTGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 13971
QY 1186 AGCC 1189
DB 13970 AGCC 13967

RESULT 14
US-09-735-934A-3/c
; Sequence 3, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
; APPLICANT: Li, Jiyin et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00851
; CURRENT APPLICATION NUMBER: US/09/735,934A
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-934A-3

Query Match      28.6%; Score 412.2; DB 4; Length 43950;
Best Local Similarity 66.0%; Pred. No. 5.6e-105;
Matches 863; Conservative 0; Mismatches 333; Indels 111; Gaps 14;

QY 2 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 61
DB 37294 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 37235
QY 62 CTCAGCTCAGCGCAACTCCG--CTCCCGGGTTCAAGGATTTCTGCTGCTGCTGCTGCTGCT 120
DB 37234 CTCAGCTCAGCGCAACTCCG--CTCCCGGGTTCAAGGATTTCTGCTGCTGCTGCTGCTGCT 37175
QY 121 CAGTACGAGTATTCAGGCAATGTCACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 37174 GAGTACGAGTATTCAGGCAATGTCACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37118
QY 181 TAGAGATGAGTATTCAGGCAATGTCACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 37117 TAGAGATGAGTATTCAGGCAATGTCACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37058
QY 241 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 37057 TGGCGGCTCAGCTCCTCCAAAGTGTGGGATTTA---TAGCATGAGCCGCTGCGCCCA 37002
QY 301 TGGCTAATTTTGTGATGAGAAACAGGTTTACATGATGTGCCAAGCTGCTGCTGCTGAGC 360
```

Db 37001 CCAGAAACCTTTAAATATGTTTGACAACTTGGGACATGCAATCTATTTCCTCAACTG 36942
QY 361 TCAGAGCACTGACCTGCTGACCTCCCAAGTCTGGGATTAACAGCGCTGACCGCTGC 420
Db 36941 TTAATACACCAAGATTTTG-----AAGACATAGTAAAGAAAACCAACAAATGTT 36892
QY 421 CTGACCTTTTATTTTATTTTATTTTAAAGACAGAGTGTCCACTCTTACCAGAGTGAAG 480
Db 36891 TCATTGAGTTTGTGTTTCTTTTAAAGACA-GGGTCTGTCTGTCTGACCCACTGGAG 36833
QY 481 TGCAGTGTGTGATACAGCTCACTGACGCTTCACTCTGAGATCAAG-CATCTCTCT 539
Db 36832 TGCAGTGTGACAAATCATATGCTCACTGACACTTCACTCTGCTCAAGCGCTTCC 36773
QY 540 GCCCTGAGCTCCCAAGTGTGGGACCAACATGACCCACTGACCTGGC----- 591
Db 36772 ATCTTGGCTCCCAAGTGTGGGATTAAGGACATGAGCCACTCCCTGGCTTAATA 36713
QY 592 ----- 591
Db 36712 TTTTATGCTGACTACATGTTGAATTAAGTCTATGCTTGGATTAACCTATATATAT 36653
QY 592 TAAATTTTATTTTATTTTATTTTATTTTGAAGACAGAGTCTC-----AACTCTGAC 644
Db 36652 ATATATATATATTTTATTTTATTTTATTTTGAAGACAGAGTCTGCTCTGCTGCTGAC 36593
QY 645 CAGGCTGAGAGTCACTGCGCAATCTTGGCTCACTGCACTCTGCTCCCGGGTTCAAG 704
Db 36592 CAGGCTGAGAGTCACTGCGCAATCTTGGCTCACTGCACTCTGCTCCCGGGTTCAAG 36533
QY 705 TTAATCTCTGCTCCAGCTCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 761
Db 36532 CCAATCTCTCTCACTCACTGCTCCCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 36473
QY 762 GCTAA--TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 818
Db 36472 GCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 36413
QY 819 TTGATCTCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 878
Db 36412 TCGATCTCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 36353
QY 879 GTGAGCCACACAGCCCGGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 938
Db 36352 GTGAGCCACACAGCCCGGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 36298
QY 939 GTTACCCAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 998
Db 36297 GTTATTCAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 36238
QY 999 CTCACGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058
Db 36237 CTCACGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 36178
QY 1059 CACCCGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1115
Db 36177 TTCCGCGCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 36118
QY 1116 TGGTCTCAACTCTGACCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1175
Db 36117 TGGTCTCAACTCTGACCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT -GAT 36060
QY 1176 TACAGCGCTGAGACACCTCAAGCGGCTAATTTATTAATAAAAA 1222
Db 36059 TACAGCGCTGAGACACCTCAAGCGGCTAATTTATTAATAATA 36013

RESULT 15
US-08-687-080-101/c
Sequence 101, Application US/08687080
Patent No. 5965427
GENERAL INFORMATION:

APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,080
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/592,126
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Shultz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 5543 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: 3' END OF INTRON 19 OF RAD50 GENOMIC
INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-101
Query Match 27.8%; Score 400.8; DB 2; Length 5543;
Best Local Similarity 72.1%; Pred. No. 3.5e-102;
Matches 610; Conservative 0; Mismatches 217; Indels 19; Gaps 6;
QY 592 TAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 651
Db 1274 TGACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1216
QY 652 GAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 711
Db 1215 GAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1156
QY 712 CCGTCCCGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768
Db 1155 CCGTCCCGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
QY 769 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 827
Db 1095 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1036
QY 828 GGACCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 887
Db 1035 TGACCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 888 CAGCGCGGCTTAT-----TTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 937
Db 975 CAGCGCGGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 916
QY 938 -TTTACCAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 996

```
Db 915 TTGTACCCAGGCTGAAGTGCATGATGTGATCTGGCTCACTGCATCTCTGCTCCCA 856
Oy 997 GGCTCAAGCGATTCCTCCCTGCTCAGCCCTCCAGCAGCTGGGATTTAGGGCACCCTGCAC 1056
Db 855 GGTTCAGAGCGATTCCTCTGCTCAGCCCTCAGAGTCTGGATTTACAGCGCGGTGCAC 796
Oy 1057 CACACCCCGCTAATTTTGTATTTTCATTTAGAGCGGGGTTTCACCAATATTGTGAGGCT 1116
Db 795 CACACCTGGCTAATTTTCTATTTTCATAGATAGATGGGTTTACATATTGGCCAGGCT 736
Oy 1117 GGCTCAAACTCTGACCTCAGGTGACCCACCCTGCTCAGCTTCGCAAGTGTGGATT 1176
Db 735 AGTCTTGAACCTCAGACCTCAGGTGATCCGCTGCTGGCTCCCAAGTGTGGGATT 676
Oy 1177 ACAGGCGTGAGCCACCTCAGCCAGCGGCTAATTTAGATAAATAATGTAGCAATGGG 1236
Db 675 ACAGCATGAGTCACTGCGCTGCGCTCTCTTTATTTGACTAGAAATCTTCAGCA 616
Oy 1237 GGGTCTTGCTATGTTGCCAGGCTGCTCAAACTTCTGCTCATGCAATCTTCCAAA 1296
Db 615 AGCATATCAGACTTCATGCAATCTTTTATACACTCT--CTCGGTTTCATTACTTT 559
Oy 1297 TGAGCCACACACCCAGCAGTCAATTTTAAACAGTTACATCTTATTAGTATAC 1356
Db 558 CTGCCCCCTATTCTACACTGCTGTTTCCCATTAATTGAATACATTTATCTTGC 499
Oy 1357 TAGAAGTAAATACATAAATGATGCAACCTGCAAAATTCAGTAGTAACAGAGTCTTTA 1416
Db 498 TCTATTGTATTAATTAAGTAAATTAATTCTGGAACAAGAGTTACAAAGTAAACTTA 439
Oy 1417 TAACTT 1422
Db 438 TACCAT 433
```

Search completed: July 20, 2003, 07:55:10
Job time : 83 secs

THIS PAGE BLANK (USPTO)